



**MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ ABDELHAMID IBN BADIS DE MOSTAGANEM**

**Faculté des Sciences Exactes et d'Informatique
Département de Mathématiques et d'Informatique
Filière Informatique**

**MEMOIRE DE FIN D'ETUDES
Pour l'Obtention du Diplôme de Master en Informatique
Option : Ingénierie des Systèmes d'Information**

THEME :

Identification faciale en 3D pour le contrôle d'accès aux
zones restreintes

Etudiantes :

- ✓ **BOUTERFA Nihad**
- ✓ **LAROUI Hadjer**

Encadrante :

- ✓ **BENTAOUZA Chahinez Mérièm**

Deuxième Année Master Ingénierie de Systèmes d'Information

Année Universitaire 2018/2019

Résumé

Pour échapper à tous les problèmes d'insécurité que l'on rencontre quotidiennement dans tous les domaines à savoir le travail, l'accès aux administrations, aux résidences... La reconnaissance faciale a été conçue. Cette technique est basée sur la capture des données faciales d'un individu pour permettre son identification à partir des données déjà enregistrées, sa performance est telle que ses recherches ne cessent de développer. Il existe deux méthodes de reconnaissance faciale la 2D qui se base uniquement sur une seule photo de l'individu puis il y a eu la reconnaissance en 3D qui est beaucoup plus performante sur laquelle nous allons nous focaliser. Ce rapport présente le développement des techniques de détection du visage et de reconnaissance faciale 3D. Les étapes principales d'un système de reconnaissance faciales 3D sont la détection du visage et son prétraitement et l'extraction caractéristiques en passant par des méthodes de classification. Pour cela, nous avons présenté, le principe des méthodes que nous avons choisies : Viola-Jones pour la détection du visage, le maillage 3D pour la reconnaissance 3D du visage et le CNN comme méthode de classification.

Mots clés : Reconnaissance Faciale, Détection du visage, Viola-Jones, Reconnaissance de visages en 3D, CNN, Maillage.

Abstract

Facial recognition was designed to solve all the problems of insecurity that we daily face everywhere in work, at home, acceding to administrations.... This technique is based on capturing the facial data of a person, these data will be compared with the already saved data. It is a such performant that so many researches are focusing in it. There are two methods of recognition the first one is based only on a single photo of the individual and the second one 3D recognition, which has more advantages than the 2D and that's what we are going focus in our chapters. These chapters show the development of facial detection and 3D facial recognition techniques. The main steps of a 3D facial recognition system are face detection, its pretreatment, extracting features with methods of classification. For this, we've described the principle of the methods we've chosen: Viola-Jones for the detection of the face and the 3D mesh for the 3D recognition of the face and the CNN as a method of classification. In this work we will implement a 3D facial recognition system which is based on the model approach.

Keywords : Facial Recognition, Face Detection, Viola-Jones, 3D Faces Recognition, CNN, Mesh.

DEDICACE

*Au nom d'Allah le Clément qui manifeste sa Miséricorde et
Louange à Lui.*

*Je dédie ce modeste travail en signe de reconnaissance à mes très
chers parents dont le soutien et les conseils étaient
indispensables.*

Ma profonde gratitude à eux.

*A mes sœurs chéries Manel et Chaïmaâ pour leur amour et leur
présence dans ma vie.*

A mes chers neveux Mohamed et Iyad.

Et enfin à toute ma famille.

Nihad,

REMERCIEMENTS

En premier lieu je remercie le très savant Dieu de m'avoir inculqué la volonté de mener à bien mon cursus universitaire.

Toute ma reconnaissance revient à mon encadrante Mme Bentaouza Chahinez Mérièm qui a eu l'amabilité de répondre à mes questions et de fournir les explications nécessaires.

J'exprime également mes remerciements au jury de la faculté de l'informatique et sciences exacte de l'Université Abd el Hamid Ibn Badis pour avoir bien voulu juger ce modeste travail.

Je remercie également mes très chers parents de m'avoir soutenue et qui ont toujours été présents ainsi que mes sœurs Manel et Chaïmaâ.

Je dois exprimer les sincère gratitude à tous les professeurs de la faculté pour leur dévouement et leur assistance tout au long de ma formation.

Je saisis cette occasion pour adresser mes profonds remerciements et mes profondes reconnaissances aux responsables et au personnel de la faculté.

Mes sincères remerciements à toutes les personnes qui par leur contribution dans ce travail, leurs conseils et leurs critiques.

DEDICACE

Je dédie ce mémoire à :

Mes parents :

Ma mère, qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.

*Mon père, qui peut être fier et trouver ici le résultat de longues années de sacrifices et de privations pour m'aider à avancer dans la vie. Puisse Dieu faire en sorte que ce travail porte son fruit ;
Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de toi.*

Mes sœurs qui n'ont cessé d'être pour moi des exemples de persévérance, de courage et de générosité.

Mon neveu Abdelhadi et ma nièce Manel à qui je porte un amour éternel.

Hadjer,

REMERCIEMENTS

Je tiens à exprimer toute ma reconnaissance à ma directrice de mémoire, Madame Chahinez Bentaouza pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter notre réflexion. Je la remercie encore de nous avoir encadré, orienté et aidé.

Je remercie mes très chers parents, Mokhtar et houaria, qui ont toujours été là pour moi.

Je remercie également mes sœurs Imene et Nabila pour leurs encouragements.

À tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

SOMMAIRE

Introduction générale.....	- 11 -
Chapitre I « Reconnaissance Faciale en générale ».....	13
I.1. Introduction.....	14
I.2. Historique	14
I.3. Reconnaissance faciale	15
I.4. Système de reconnaissance faciale.....	15
I.5. Architecture générale d'un système de reconnaissance faciale	15
I.5.1 Détection de visage	15
I.5.2. Extraction de caractéristiques du visage	17
I.5.3 Comparaison des caractéristiques (classification).....	17
I.7. Principales difficultés de la reconnaissance faciale.....	17
I.8. Les principaux avantages de la reconnaissance faciale	20
I.9. Conclusion	20
Chapitre II « Identification Faciale en 3D »	21
II.1. Introduction	22
II.2. Définition.....	22
II.3. Système d'acquisition 3D.....	22
II.3.1. Vision active	22
II.3.2. Vision passive	22
II.4. Processus de la reconnaissance de visages 3D.....	23
II.4.1. Les approches	23
II.5. Avantage.....	25
II.6. Défis de la reconnaissance faciale 3D	25
II.7. Classification des visages 3D	26
II.7.1. Méthodes de classification.....	27
II.8. Conclusion.....	28
Chapitre III « Les Méthodes Utilisées »	30
III.1. Introduction	31
III.2. Détection du visage	31
III.3. Extraction des points caractéristiques 2D et 3D	36
III.4. Classification des visages 3D.....	38

III.4.1. L'apprentissage en profondeur.....	39
III.4.2. Couche de convolution.....	39
III.4.3. Couche de sous-échantillonnage (Pooling)	40
III.5 Conclusion	41
Chapitre IV « La Conception et La Réalisation »	42
IV.1. Introduction.....	43
IV.2. Environnement matériel et logiciel.....	43
IV.2.1. Ressources utilisées.....	43
IV.2.2. Langage de programmation.....	43
IV.3. Système « FaceIt ».....	44
IV.3.1. Méthodes de reconnaissance faciale 3D.....	45
IV.4. Organigramme du système de reconnaissance du visage en 3D	46
IV.5. Présentation de l'application	47
IV.5.1. Fonctionnement du système	47
IV.5.2. Capture	48
IV.5.3. Interface principale.....	48
IV.5.4. Capture	49
IV.5.5. Détection du visage	49
IV.5.6. Extraction des caractéristiques 2D et 3D	49
IV.5.7. Fichier JSON	50
IV.5.8. Apprentissage.....	51
IV.5.9. Classification.....	51
IV.5.10. Ajout d'un nouveau visage.....	52
IV.6. Conclusion	52
Conclusion générale	53
REFERENCES ET WEBOGRAPHIE.....	54

Liste des figures

<u>Figure I.1. Processus de reconnaissance faciale</u>	11
<u>Figure I.1. Schéma général d'un système de reconnaissance faciale</u>	15
<u>Figure I. 2. Variation de luminosité</u>	18
<u>Figure I.3. Variation de poses</u>	19
<u>Figure I.4. Variation des expressions</u>	19
<u>Figure II. 1. Etapes d'un système de reconnaissance</u>	23
<u>Figure II. 2. Les données manquantes dues aux auto-occultation et à l'absorption du rayon laser par des zones sombres du visage</u>	25
<u>Figure II. 3. Déformations sous les expressions du visage (a). Les données manquantes dues à l'occlusion d'auto (b).</u>	26
<u>Figure III.1. Exemple de détection de visage</u>	31
<u>Figure III.2. Caractéristiques de base de type rectangle Haar</u>	32
<u>Figure III.3. La représentation d'une image intégrale</u>	33
<u>Figure III.4. Caractéristiques pseudo-haar avec différentes orientations</u>	34
<u>Figure III.5. Architecture de la cascade</u>	35
<u>Figure III.6. Exemple : (a) une image du visage originale, (b) sa projection (SOM-Face) et (c) image reconstruite</u>	36
<u>Figure III.7. (a) Image texture (b) Image 2.5D (c) Image 3D</u>	36
<u>Figure III.8. De 2D vers la 3D</u>	36
<u>Figure III.9. L'architecture d'un réseau de neurone convolutif</u>	38
<u>Figure III.10. L'opération de convolution</u>	39
<u>Figure III.11. L'opération de sous échantillonnage</u>	39
<u>Figure IV.1. Interface de PyQT4</u>	43
<u>Figure IV.2. Organigramme de système de reconnaissance du visage en 3D</u>	45
<u>Figure IV.3. Schéma du système « FaceIt »</u>	46
<u>Figure IV.4. Caméra utilisée</u>	47
<u>Figure IV.5. Interface principale</u>	47
<u>Figure IV.6. Détection du visage</u>	48
<u>Figure IV.7. Visage détecté</u>	48
<u>Figure IV.8. Extraction des caractéristiques faciales</u>	49
<u>Figure IV.9. Fichier JSON</u>	49
<u>Figure IV.10. Résultat d'apprentissage</u>	50
<u>Figure IV.11. Classification de l'individu</u>	50
<u>Figure IV.12. Ajout d'une nouvelle personne</u>	51

Liste des abréviations

1D : Dimension (dans l'espace).

2D : Dimensions (dans l'espace).

2.5D : Dimensions (dans l'espace).

3D : Dimensions (dans l'espace).

ACP : Principal Component Analysis, Analyses Des Composantes Principales.

Adaboost : Adaptation Boosting.

CNN : Réseaux de neurones convolutionnel.

FRVT : Recognition Vendor Test.

ICP : Independent Component Analysis, Analyses en Composantes Indépendantes.

ILSVRC : ImageNet Large Scale Visual Recognition Challenge, Compétition ImageNet de Reconnaissance Visuelle à Grande Échelle.

JSON : JavaScript Object Notation.

Licence BSD : Berkeley Software Distribution.

OpenCV : Open Source Computer Version, Bibliothèque graphique libre.

OS : Operation System.

Qt Linguist : Quick Tool Linguist.

SVM : Machine à Vecteurs de Support.

XML : Extensible Markup Language.

Introduction générale

Savoir déterminer de manière à la fois efficace et exacte l'identité d'un individu est devenu un problème critique car aujourd'hui, l'accès sécurisé et la surveillance constituent un sujet primordial, aujourd'hui un individu a besoin d'être identifié dans de tous les lieux et situations qui requièrent une sécurité (par exemple pour pénétrer dans son immeuble ou ouvrir la porte de son appartement, accéder à son compte bancaire, à son réseau social...) [1].

La méthode la plus récente d'identification et que nous allons aborder dans nos chapitres, est la reconnaissance faciale, qui a remplacé toutes les méthodes traditionnelles telles que les mots de passe, les badges, les clefs... ces dernières qui non seulement nécessitent la mémorisation, mais ne nous protègent pas d'éventuels vols et d'insécurité. Ce sont ces inconvénients-là qui ont abouti à l'idée d'exploiter les informations biométriques à la place (empreintes, données faciales...) comme moyen d'identification, notamment la reconnaissance faciale qui, en dehors de tous les problèmes cités, nous facilite la tâche.

Le présent mémoire comporte trois chapitres.

Dans le premier chapitre, intitulé « Reconnaissance Faciale En Général », nous allons montrer un aperçu sur l'histoire de naissance de la reconnaissance faciale ainsi que sur les améliorations qu'elle a subies. Nous allons décortiquer chacune de ses étapes.

En effet, pour concevoir une machine arrivant à réaliser la réception de l'image puis son traitement, chose qui se fait naturellement par notre rétine et les centres nerveux d'intégration, est une tâche assez difficile et nécessite un processus assez complexe comme le montre la Figure II.1.

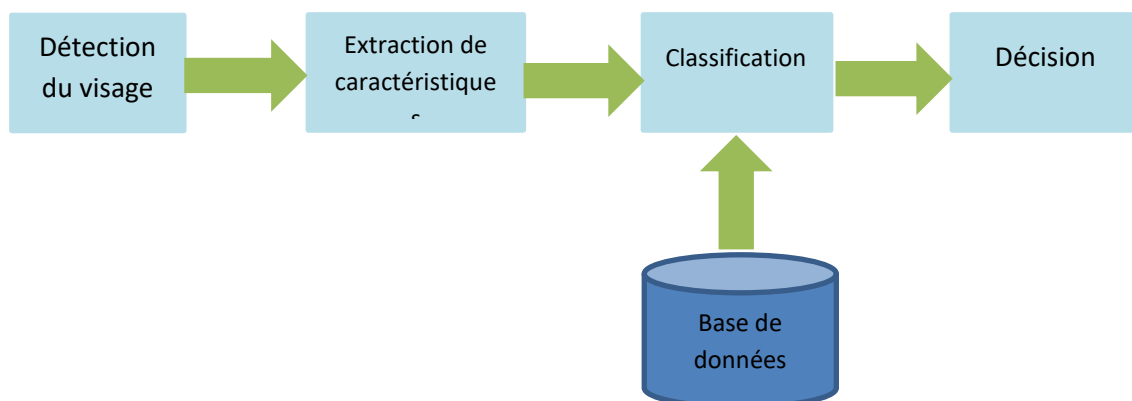


Figure.I.1. Processus de reconnaissance faciale [2].

[Introduction générale]

Depuis cette percée, ce domaine est devenu le centre des recherches, toutes les techniques utilisées ont connues plusieurs améliorations pour arriver à de meilleurs résultats.

Dans le second chapitre, nous allons nous focaliser sur la reconnaissance faciale en 3D qui est une méthode très prometteuse car elle permet de résoudre les problèmes qui sont dus aux changements touchant l'endroit (éclairage variable) et les caractéristiques dynamiques spécifiques à chaque personne, comme le mouvement de la tête, l'évolution de la pose ou expressions faciales.

Et donc facilite la tâche de reconnaissance. Cette dernière permet en plus, de faire des représentations meilleures, telles que des modèles 3D de visages ou des images de résolutions augmentées, peuvent être obtenus à partir des séquences vidéo et être employés pour améliorer les performances des systèmes. Enfin, la reconnaissance faciale en 3D permet d'apprendre ou de mettre à jour les modèles dans le temps.

Dans le troisième chapitre, nous allons discuter des trois méthodes que nous avons choisies parmi toutes les méthodes de reconnaissance 3D existantes.

Dans le quatrième et dernier chapitre, nous présenterons la conception et la réalisation de notre projet.

Enfin nous allons clôturer notre travail par une conclusion générale résumant tous ce qui a été mentionné.

Chapitre I

Reconnaissance Faciale en générale

[Chapitre I : Reconnaissance faciale en général]

I.1. Introduction

Partout où on va et dans tous les domaines, c'est la sécurité qui pose un problème, les méthodes classiques telles qu'un mot de passe, des pièces d'identités, des clés... peuvent ne pas garantir une sécurité absolue et être inefficace. En effet, un mot de passe peut être oublié, trouvé par une autre personne, les clés peuvent être volées ...

Donc la probabilité d'échapper à un problème quelconque utilisant ces méthodes n'est jamais nulle et la sécurité assurée dans ce cas-là est très faible, c'est alors que survint un concept beaucoup plus moderne qui a pu booster les performances de sécurité, c'est bien la reconnaissance faciale. Inutile de toujours faire attention à ce que l'on possède ou d'avoir nécessairement une bonne mémoire pour apprendre ses codes par cœur, une simple identification des traits physiques des individus suffit pour assurer la surveillance.

I.2. Historique

Le premier système semi-automatisé de la reconnaissance faciale a été innové dans les années 1960, il est basé sur les calculs des distances entre les yeux, les oreilles, le nez et la bouche après les avoir localisés sur la photo et de les comparer par la suite aux distances de référence [3].

Avec le temps, ce système a subi plusieurs améliorations, en effet, Goldstein et Lesk ont pu définir 21 marqueurs spécifiques vers les années 1970, tels que la couleur des cheveux et l'épaisseur de la lèvre pour automatiser la reconnaissance. Néanmoins, les mesures et les emplacements ont été calculés manuellement [4].

Ensuite, en 1988, Kirby et Sirovich ont appliqué l'analyse en composantes principales (ACP), une technique standard de l'algèbre linéaire. Cela demeure une étape importante vu qu'elle montre la nécessité de certaines valeurs pour pouvoir coder d'une manière convenable, une image alignée et normalisée [5].

En 1991, Turk et Pentland ont découvert que l'erreur résiduelle dans la technique ACP, peut être utilisée pour détecter un visage dans une image, une découverte qui a permis la reconnaissance faciale automatique en temps réel. Bien que l'approche fût quelque peu limitée par des facteurs environnementaux, elle a néanmoins créé un intérêt significatif pour promouvoir le développement des technologies de la reconnaissance faciale automatique. Cette technologie a été mise en essai en janvier 2001 lors de la finale du championnat de football américain SUPER BOWL en capturant des images de surveillance puis comparées à une base de données numérique [6].

[Chapitre I : Reconnaissance faciale en général]

I.3. Reconnaissance faciale

La reconnaissance faciale est une technique d'identification biométrique permettant d'identifier un individu. Cela est courant dans le domaine de la sécurité, afin d'identifier des criminels ou des disparus et autoriser l'accès à certains lieux publics [7].

Il existe deux méthodes de reconnaissance faciale, l'une en 2D qui se base uniquement sur une seule photo de l'individu, l'autre, en 3D qui est une modernisation de la première et c'est ce que nous allons voir en détails dans le prochain chapitre.

I.4. Système de reconnaissance faciale

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 [8].

I.5. Architecture générale d'un système de reconnaissance faciale

Un système de reconnaissance faciale peut être structuré comme le montre la figure I.1.

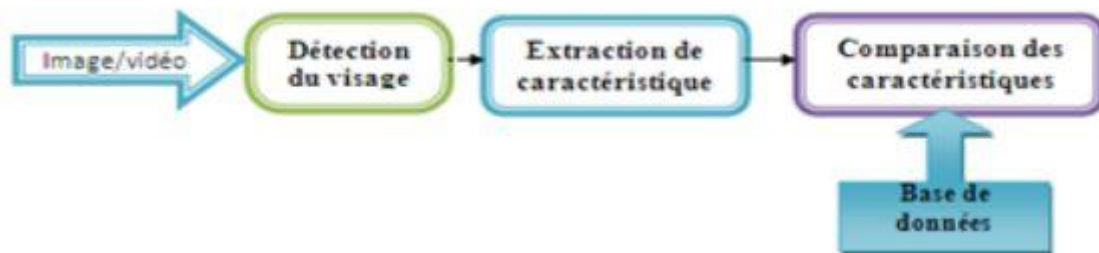


Figure I.1. Schéma général d'un système de reconnaissance faciale [28].

Le processus de la reconnaissance faciale de visage s'effectue en trois étapes principales :

I.5.1 Détection de visage

Cette étape consiste à l'extraire de l'image de visage, 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 [28].

Les méthodes de détection de visages peuvent être subdivisées en quatre catégories : Son efficacité a une influence directe sur les performances du système de reconnaissance de visages.

[Chapitre I : Reconnaissance faciale en général]

Il existe plusieurs méthodes pour la détection de visages, certaines utilisent la couleur de la peau, la forme de la tête, l'apparence faciale, alors que d'autres combinent plusieurs de ces caractéristiques [10].

Les méthodes de détection de visages peuvent être subdivisées en quatre catégories :

➤ **Approches basées sur les connaissances acquises**

Ces méthodes se basent sur la connaissance des différents éléments qui constituent un visage et des relations qui existent entre eux. Ainsi, les positions relatives de différents éléments clés tels que la bouche, le nez et les yeux sont mesurées pour servir ensuite à la classification visage ou non visage. Le problème dans ce type de méthode est qu'il est difficile de bien définir de manière unique un visage. Si la définition est trop détaillée, certains visages seront ratés tandis que si la description est trop générale, le taux de faux positifs montera en flèche [10].

➤ **Approches basées sur le « Template-matching »**

La détection des visages se fait à travers un apprentissage d'exemples standards de visages ou d'images frontales contenant des visages. La procédure se fait en corrélant les images d'entrées et les exemples enregistrés (gabarits) et le résultat donne la décision finale soit de l'existence ou non d'un visage [11].

➤ **Approches basées sur l'apparence (apprentissage)**

La différence entre cette méthode et "Template matching" est que les modèles (Template) sont lus à partir des images d'apprentissage qui doivent être représentatives et faites à différentes positions du visage. Généralement l'approche basée sur l'apparence se base sur des techniques d'analyse statistiques (pourcentage d'existence des modèles dans l'image) et d'apprentissage automatique pour trouver des caractéristiques significatives des visages et des non visages [12].

On peut citer parmi celles-ci, la méthode basée sur les réseaux de neurones de Rowley et al, la méthode de Schneiderman et Kanade basée sur un classifieur de Bayes naïf ainsi que le fameux algorithme de Viola et Jones fonctionnant en temps réel [13].

➤ **Approches basées sur des caractéristique invariantes**

Ces approches sont utilisées principalement pour la localisation de visage. Les algorithmes développés visent à trouver les caractéristiques structurales existantes même si la pose, le point de vue, ou la condition d'éclairage changent. Puis ils emploient ces caractéristiques invariables pour localiser les visages. Nous pouvons citer deux familles de méthodes appartenant à cette approche : Les méthodes basées sur la couleur de la peau, et les méthodes basées sur les caractéristiques de visage, elles consistent à localiser les cinq caractéristiques (deux yeux, deux narines, et la jonction nez/lèvre) pour décrire un visage typique [13].

[Chapitre I : Reconnaissance faciale en général]

I.5.2. Extraction de caractéristiques du visage

Après la détection de la zone de visage, on doit extraire les informations utiles qui reviennent à établir un modèle du visage (vecteur de caractéristiques telles que les yeux, le nez, la bouche...). Ces informations nécessaires pour que le visage d'une personne ne ressemble pas à celui d'une autre personne et en même temps qu'il ressemble à lui-même dans d'autres conditions d'acquisition. C'est une étape prétraitement nécessaire dont l'efficacité a une influence directe sur la performance du système de reconnaissance de visage [28].

On peut distinguer deux pratiques différentes :

- **Les pratiques globales** : repose sur l'extraction de régions entières du visage, elle est souvent implémentée avec une approche globale de reconnaissance de visage [14]. L'une des méthodes la plus largement utilisée pour la représentation du visage dans son ensemble est ACP [28].
- **Les pratiques locales** : 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 [28]. Elles sont utilisées avec une méthode locale de reconnaissance et aussi pour l'estimation de la pose du visage [14].

Par ailleurs, plusieurs études ont été menées afin de déterminer les caractéristiques qui semblent pertinentes pour la perception, la mémorisation et la reconnaissance d'un visage humain. Par exemple, les caractéristiques pertinentes rapportées sont : les cheveux, le contour du visage, les yeux et la bouche. Cette étude a également démontré le rôle important que joue le nez dans la reconnaissance faciale à partir des images de profil. En effet, dans ce cas de figure, il est évident que la forme distinctive du nez est plus intéressante que les yeux ou la bouche [14].

I.5.3 Comparaison des caractéristiques (classification)

Selon les caractéristiques extraites précédemment, il faut introduire une mesure de similarité afin d'estimer la différence entre deux visages pour identifier une personne [28]. Cela se fait en comparant un modèle du demandeur (une ou plusieurs images de test), avec un modèle stocké (une ou plusieurs images de référence) [15]. Plusieurs méthodes sont utilisées pour la comparaison de visages [28].

On trouve dans la littérature : calcul de distance, calcul de similarité. D'autres méthodes se basent sur la classification des caractéristiques par un seul classificateur (SVM, classificateur bayésien, etc.) ou par plusieurs (Adaboost) [28].

I.7. Principales difficultés de la reconnaissance faciale

Bien que l'idée de la reconnaissance faciale soit concrétisée et appliquée dans notre vie quotidienne, la tâche de la construction d'un système automatique d'identification idéal représente un sérieux défi vu les nombreux problèmes rencontrés. Ce défi est d'autant plus

[Chapitre I : Reconnaissance faciale en général]

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 [16].

Ces variations sont recensées comme suit :

- **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 I.2). 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 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 [16].



Figure I.2. Variation de luminosité [16].

- **Variation de la pose**

Le taux de reconnaissance de visage baisse considérablement quand des variations de pose sont présentes dans les images [17].

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 I.3.) [17].



Figure I.3. Variation de poses [17].

- **Expressions faciales**

Un autre facteur qui affecte l'apparence du visage est l'expression faciale (voir figure I.4.) 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 [17].

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 [17].



Figure I.4. Variation des expressions [17].

- **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 [17].

[Chapitre I : Reconnaissance faciale en général]

- **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 [17].

- **Vieillessement 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 [17].

I.8. Les principaux avantages de la reconnaissance faciale

- Cette technologie est peu intrusive, et peut même se dérouler à l'insu de l'utilisateur.
- Certains de ces systèmes sont aujourd'hui capables de reconnaître des individus même avec des subterfuges (fausses moustaches, barbe, lunettes...).
- Elle ne demande aucune action de l'utilisateur.
- Il n'y a pas de contact physique.
- Cette technique est peu coûteuse [18].

I.9. Conclusion

Nous avons présenté dans ce chapitre ce qu'est la reconnaissance faciale globalement en citant les techniques de cette dernière. La difficulté éprouvée quand 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. Dans le deuxième chapitre nous allons nous focaliser sur la reconnaissance faciale en 3D.

Chapitre II

Identification Faciale en 3D

[Chapitre II : Identification faciale en 3D]

II.1. Introduction

L'évolution de la technologie de nos temps actuels, a non seulement pu remplacer toutes les manières traditionnelles et classiques, à savoir une feuille et un stylo, des codes d'accès... par la reconnaissance faciale mais a pu subir aussi beaucoup de modifications innovantes pour nous aboutir à de meilleurs résultats de sécurité et de l'identification.

II.2. Définition

La reconnaissance de visages en 3D constitue une modalité prometteuse qui permet de résoudre les problèmes liés aux différents facteurs de variabilité, notamment l'éclairage, la pose, l'occultation... [19].

L'avantage principal des approches basées sur le modèle 3D, et ce qui fait la différence entre la représentation en 2D et celle-ci, réside dans le fait que le modèle 3D conserve toutes les informations sur la géométrie du visage, ce qui permet d'avoir une représentation réelle de ce dernier. Plusieurs méthodes ont été proposées afin de permettre la reconnaissance de visages en 3D [19].

II.3. Système d'acquisition 3D

Au cours de ces 30 dernières années, le développement de systèmes d'acquisition de données 3D a fortement progressé. Ces systèmes de mesure sont apparus dans les laboratoires dans les années 80 et s'implantent de nos jours de plus en plus rapidement et de plus en plus massivement dans de nombreux secteurs [20].

En 1988, Besl présente un état de l'art des méthodes d'acquisitions 3D actives et optiques et compare les différents systèmes existant à l'époque [20].

En 2000, Beraldin décrit les principes et les technologies des systèmes d'acquisition 3D actifs. Certains de ces différents systèmes d'acquisition utilisés pour la reconstruction 3D de visages utilisent des systèmes de vision actifs, d'autres utilisent des systèmes de visions passifs [20].

II.3.1. Vision active

Elle utilise un capteur actif illuminant les objets de la scène en projetant sur leur surface des motifs structurés (un spot, un plan lumineux...), ce qui permet une indépendance des conditions d'éclairage de l'environnement et minimise la complexité des traitements à effectuer pour reconstruire l'image (le scanner 3D équipé d'un laser, en est un exemple) [21].

II.3.2. Vision passive

Appelée vision binoculaire ou vision stéréoscopique, elle utilise deux caméras fixes pour générer les données 3D des visages. La vision stéréoscopique se veut la plus proche possible de la stratégie perceptive humaine, plaçant les caméras sur le même axe de vision dans une

[Chapitre II : Identification faciale en 3D]

configuration convergente. Chaque image droite et gauche représente la même scène vue sous un angle différent [21].

II.4. Processus de la reconnaissance de visages 3D

Le schéma classique d'un processus de reconnaissance d'objets, présenté dans la figure II.1, distingue deux étapes : la première est l'apprentissage qui consiste à apprendre la description des objets à partir d'une base d'exemples (galerie) pour construire des classes, la seconde est la classification qui reconnaît un objet inconnu (requête) à partir de sa description en lui assignant une classe [22].

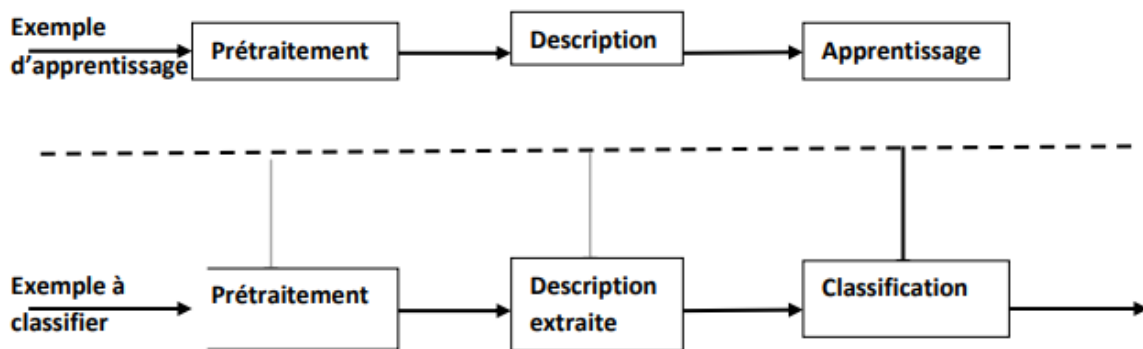


Figure II. 1. Etapes d'un système de reconnaissance [22].

II.4.1. Les approches

Les approches existantes qui s'adressent au problème de la reconnaissance du visage 3D peuvent être classées selon trois catégories d'approches : Approches basées sur le modèle, Approches 3D, elles-mêmes subdivisées en approches locales et globales, Approches 3D+2D regroupées en trois : approches 3D+2D globales, locales et hybrides.

➤ Approches basées sur le modèle

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 [23].

➤ Approches 3D

Les approches 3D sont subdivisées en deux catégories : les approches locales qui utilisent la géométrie de la surface du visage et les approches globales 3D (holistiques).

• Approches 3D globales (holistiques)

Pour l'appariement *holistique*, un algorithme basé sur l'ACP (analyse en composantes principales) est utilisé [22].

L'Analyse en Composantes Principales est l'une des méthodes d'analyse de données multivariées les plus utilisées [24].

[Chapitre II : Identification faciale en 3D]

Nous pouvons subdiviser les méthodes holistiques 3D en deux catégories : approches surfaciques et approches volumiques.

Les techniques de l'analyse en composantes principales ont été largement utilisées dans la reconnaissance faciale 2D. Plus récemment, ces techniques ont été aussi étendues aux données 3D de visage. On prend l'exemple de Heshner et al qui ont appliqué l'analyse en composantes principales sur une base de données contenant 222 images de profondeur (37 personnes), avec six expressions faciales différentes [19].

- **Approches 3D locales (ou géométriques)**

La localité consiste à segmenter le visage en régions qui sont généralement associées avec un algorithme d'ICP (Iterative Closest Point) qui est utilisé pour l'alignement, ce qui minimise une distance quadratique carrée entre les points les plus proches dans un processus itératif [25].

L'introduction de la géométrie 3D du visage dans la reconnaissance de personnes est considérée comme une voie très prometteuse pour améliorer sa performance et pour la robustesse de la reconnaissance face aux variations d'expressions. Cependant, l'un des défis ouverts est celui de la robustesse de ces approches 3D aux expressions faciales [26].

- **Approches Multimodales 3D+2D**

En effet, il s'agit de techniques qui combinent des données 2D et 3D sur le visage pour améliorer les performances et la robustesse de la reconnaissance. Récemment, plusieurs approches basées sur ce principe ont été développées. Cette technique requiert le maillage polygonal comme méthode de représentation 3D [27].

- **Approches 2D/3D globales**

Combinant deux méthodes de localisation pour extraire les points 3D et les points de texture 2D pour couvrir l'information de tous les pixels de l'image [19].

L'idée est de projeter les données initiales dans un sous-espace qui représente les données de manière optimale selon un critère précis : par exemple, dans le cas de l'ACP (analyse en composantes principales) la variance est minimisée, dans l'ACI (Analyse en Composantes Indépendantes), l'indépendance des composantes est visée [19].

- **Approches 2D/3D locales**

L'une des approches les plus populaires pour représenter la surface est celle basée sur les descripteurs locaux ou une signature qui décrit la région locale. Les positions de la surface où sont estimés les descripteurs locaux peuvent être simplement choisis d'une façon exhaustive ou aléatoire. Dans le cas d'une sélection exhaustive, la redondance dans des régions, ayant une légère variation de la forme, rend cette technique inefficace [19].

La sélection aléatoire peut négliger certaines zones ayant une structure géométrique distincte ce qui réduit la performance de l'algorithme. Ainsi, extraire un ensemble représentatif de points primitifs permet de remédier à ces problèmes [19].

[Chapitre II : Identification faciale en 3D]

- **Une approche hybride**

En 2006, Mian, et al. proposent un algorithme de reconnaissance hybride: locale et globale. La localité consiste à segmenter le visage en trois régions qui sont appariées avec un algorithme d'ICP (Iterative Closest Point) modifié [22].

L'ICP (Iterative Closest Point) est un algorithme utilisé pour l'alignement en minimisant une distance quadratique carrée entre les points les plus proches dans un processus itératif. Pour l'appariement holistique un algorithme basé sur l'ACP est utilisé [22].

II.5. Avantage

- **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 [28].

- **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 [28].

- **Unicité des images 3D**

La quantité de points de mesure et de données biométriques saisies est telle qu'elle pourrait distinguer des jumeaux monozygotes [28].

II.6. Défis de la reconnaissance faciale 3D

Dans des conditions non contrôlées, les scans 3D ont souvent des données manquantes à cause de l'auto-occultation ou à l'absorption du rayon laser par des zones sombres du visage (cheveux ou les sourcils), ce qui peut générer des données manquantes dans le maillage 3D. Comme illustré dans la Figure II.6 [20].



Figure II. 2. Les données manquantes dues aux auto-occultation et à l'absorption du rayon laser par des zones sombres du visage [20].

Les premières images de la figure montrent des auto-occultations tandis que les deux dernières montrent des occultations suite à la main et au lunettes portées par l'individu lors du balayage. En outre, les variations dans les données des visages aux expressions faciales peuvent provoquer également des déformations du maillage 3D. Figure.2.3 illustre visages expressifs de la rangée du bas (comme la maille 3D texturé). La rangée du haut montre le maillage résultant 3D avec des déformations [29].

Toute approche 3D du visage de reconnaissance doit correspondre avec succès aux scans visage en présence d'une expression à base de déformations et / ou moins de données manquantes (comme illustré dans la Figure. 2.3.a respectivement et Figure II.6.2 b à la bonne qualité, neutre, modèle frontal 3D [28].

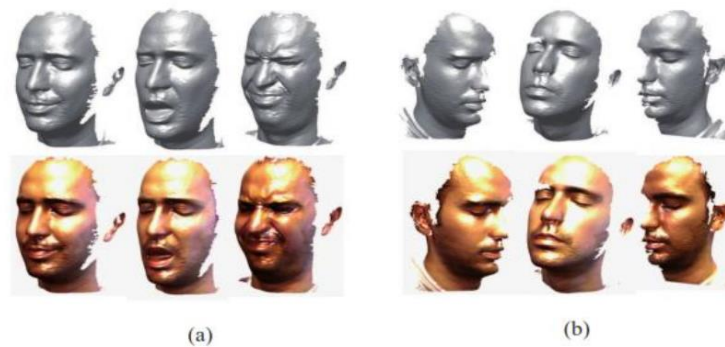


Figure II. 3. Déformations sous les expressions du visage (a). Les données manquantes dues à l'occlusion d'auto (b) [29].

II.7. Classification des visages 3D

La classification automatique des images consiste à attribuer automatiquement une classe à une image à l'aide d'un système de classification. On retrouve ainsi la classification d'objets, de scènes, de textures, la reconnaissance de visages et de caractères [42].

[Chapitre II : Identification faciale en 3D]

Il existe deux principaux types d'apprentissage : l'apprentissage supervisé et l'apprentissage non-supervisé. Dans l'approche supervisée, chaque image est associée à une étiquette qui décrit sa classe d'appartenance. Dans l'approche non supervisée où on ne dispose d'aucun échantillon d'apprentissage préalable [42].

II.7.1. Méthodes de classification

De nombreuses méthodes classiques ont été consacrées, elles peuvent être séparées en deux grandes catégories : les méthodes de classification supervisée et les méthodes de classification non supervisée [42].

➤ Méthodes supervisées

L'objectif de la classification supervisée est principalement de définir des règles permettant de classer des objets dans des classes à partir de variables qualitatives ou quantitatives caractérisant ces objets. On dispose au départ d'un échantillon dit d'apprentissage dont le classement est connu. Cet échantillon est utilisé pour l'apprentissage des règles de classement. Il est nécessaire d'étudier la fiabilité de ces règles pour les comparer et les appliquer, évaluer les cas de sous apprentissage ou de sur apprentissage (complexité du modèle). On utilise souvent un deuxième échantillon indépendant, dit de validation ou de test [42].

➤ Méthodes non supervisées

Procède de la façon contraire. C'est à dire ne nécessitent aucun apprentissage et aucune tâche préalable d'étiquetage manuel. Elle consiste à représenter un nuage des points d'un espace quelconque en un ensemble de groupes appelé Cluster.

Il lié généralement au domaine de l'analyse des données comme l'ACP. Un «Cluster» est une collection d'objets qui sont «similaires» entre eux et qui sont «dissemblables » par rapport aux objets appartenant à d'autres groupes [42].

• CNN

Les réseaux de neurones convolutifs ont une méthodologie similaire à celle des méthodes traditionnelles d'apprentissage supervisé : ils reçoivent des images en entrée, détectent les features de chacune d'entre elles, puis entraînent un classifieur dessus. L'architecture spécifique du réseau permet d'extraire des features de différentes complexités, des plus simples au plus sophistiquées. L'extraction et la hiérarchisation automatiques des features, qui s'adaptent au problème donné, constituent une des forces des réseaux de neurones convolutifs : plus besoin d'implémenter un algorithme d'extraction "à la main", Contrairement aux techniques d'apprentissage supervisé, les réseaux de neurones convolutifs apprennent les features de chaque image. Les réseaux font tout le boulot d'extraction de features automatiquement, contrairement aux techniques d'apprentissage

[Chapitre II : Identification faciale en 3D]

•SVM

Afin d'accomplir la classification des visages spécifiques des individus [] et d'aborder des problèmes très divers comme le classement, la régression, la fusion, etc... Il existe des machines à vecteurs de support, ou séparateurs à vaste marge SVM, qui est une technique récente d'apprentissage statistique, proposée par V. Vapnik. [8].

Séparateurs à vaste marge (SVM), classifieur de Bayes et Adaboost ont été examinés dans l'étude de Guo et Dyer pour le problème de classification faciale des expressions. Trente-quatre points ont été sélectionnés manuellement dans chaque image du visage puis un filtre de Gabor a été utilisé aux points sélectionnés pour extraire les caractéristiques [8].

•ACP

L'ACP est né des travaux de M.A. Turk et A.P. Pentland au MIT Media Lab, en 1991 [3]. Il est aussi connu sous le nom de Eigenfaces car il utilise des vecteurs propres et des valeurs propres [30].

Outre le fait de réduire les dimensions de l'espace original des variables, l'ACP peut aussi être utilisée comme une technique de classification, et ainsi les relations entre les variables et individus peuvent être mises en évidence [43].

Cet algorithme s'appuie sur des propriétés statistiques bien connues et utilise l'algèbre linéaire. Il est relativement rapide à mettre en œuvre mais il reste sensible aux problèmes d'éclairage, de pose et d'expression faciale [30].

L'idée principale consiste à exprimer les M images d'apprentissage selon une base de vecteurs orthogonaux particuliers, contenant des informations indépendantes d'un vecteur à l'autre. Ces nouvelles données sont donc exprimées d'une manière plus appropriée à la reconnaissance du visage [30].

Nous voulons extraire l'information caractéristique d'une image de visage, pour l'encoder aussi efficacement que possible afin de la comparer à une base de données de modèles encodés de manière similaire. En termes mathématiques, cela revient à trouver les vecteurs propres de la matrice de covariance formée par les différentes images de notre base d'apprentissage [30].

II.8. Conclusion

Dans ce chapitre, nous avons passé en revue les principales techniques de reconnaissance de visages 3D et les différentes approches utilisées. Il est à noter que les méthodes basées sur les caractéristiques faciales en 3D sont robustes aux variations de la pose, ainsi qu'à l'échelle, changement d'éclairage, et parfois les expressions faciales. Nous avons mentionné les défis de ces méthodes à savoir, les occultations, les absorptions des rayons laser dans des conditions non contrôlées. Nous avons cité les méthodes de classification les plus connues et dont la plus fiable est la CNN que nous allons détailler dans le troisième chapitre.

[Chapitre II : Identification faciale en 3D]

Les approches locales ont prouvé leur efficacité dans le cas de la présence des expressions faciales, mais elles souffrent d'un manque de robustesse aux occultations. Enfin, les approches hybrides semblent intéressantes pour résoudre ce problème. Ces techniques sont très prometteuses et ne cessent de se développer. Toute approche est basée sur une ou plusieurs méthodes précises, dans le troisième chapitre nous allons nous focaliser sur les méthodes que nous allons choisir.

Chapitre III

Les Méthodes Utilisées

III.1. Introduction

Comme nous l'avons déjà cité dans les chapitres précédents la détection et l'identification du visage reste une technique imparfaite du fait des variations du visage et de la scène.

Pour aboutir à une reconnaissance faciale 3D meilleure plusieurs méthodes ont été proposées et pour la détection du visage et pour la reconnaissance faciale ainsi que pour la classification. Dans ce chapitre nous allons détailler les méthodes que nous avons choisies.

III.2. Détection du visage

- **Avec la méthode de viola et Jones**

- **Définition**

La méthode de Viola et Jones est une méthode de détection d'objet dans une image numérique, elle fait partie des toutes premières méthodes capables de détecter efficacement et en temps réel des objets dans une image. Inventée à l'origine pour détecter des visages, elle peut également être utilisée pour détecter d'autres types d'objets comme des voitures ou des avions [30].

En tant que procédé d'apprentissage supervisé, la méthode de Viola et Jones nécessite de quelques centaines à plusieurs milliers d'exemples de l'objet que l'on souhaite détecter, pour entraîner un classifieur. Une fois son apprentissage réalisé, ce classifieur est utilisé pour détecter la présence éventuelle de l'objet dans une image en parcourant celle-ci de manière exhaustive, à toutes les positions et dans toutes les tailles possibles [30].

- **Fonctionnement**

La méthode de viola et Jones est une approche basée sur l'apparence, elle permet d'analyser une image, sans avoir besoin d'étudier chaque pixel la constituant. En effet, la notion d'image intégrale permet de définir plusieurs zones rectangulaires au sein d'une image. L'intérêt de cette technique réside dans le fait qu'elle offre la possibilité d'accéder à la valeur des autres zones à gauche et au-dessus de la zone sur laquelle nous sommes [31]. Cette méthode bénéficie d'une implémentation sous licence BSD dans OpenCV [8].

Viola-Jones comporte 4 étapes principales, et vous en apprendrez plus sur chacune d'elles dans les sections suivantes [34] :

- Sélection de caractéristiques de type Haar
- Créer une image intégrale
- Suivre une formation AdaBoost
- Création de cascades de classificateurs

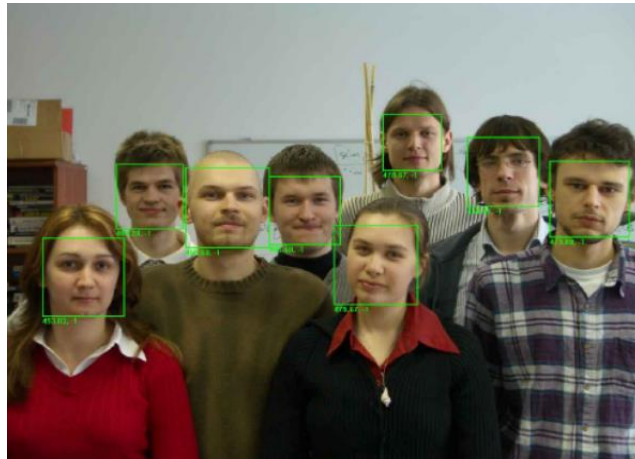


Figure III.1. Exemple de détection de visage [30]

- **Features (caractéristiques)**

Une feature est une information dans une image qui est pertinente pour résoudre un problème donné. Cela peut être quelque chose d'aussi simple qu'une valeur de pixel unique ou plus complexe comme des bords, des angles et des formes. Vous pouvez combiner plusieurs features simples en une feature complexe [34].

L'application de certaines opérations à une image produit des informations qui pourraient également être considérées comme des features. La vision par ordinateur et le traitement des images disposent d'un grand nombre de features utiles et d'opérations d'extraction de features [34].

Fondamentalement, toute propriété inhérente ou dérivée d'une image peut être utilisée comme une feature pour résoudre des tâches [34].

- **Sélection de caractéristiques de type Haar**

Tous les visages humains partagent certaines similitudes. Si vous regardez une photo montrant le visage d'une personne, vous verrez, par exemple, que la région des yeux est plus sombre que l'arête du nez. Les joues sont également plus brillantes que la région des yeux.

Nous pouvons utiliser ces propriétés pour nous aider à comprendre si une image contient un visage humain [34].

Un moyen simple de déterminer quelle région est la plus claire ou la plus sombre consiste à résumer les valeurs de pixels des deux régions et à les comparer. La somme des valeurs de pixels dans la région la plus sombre sera inférieure à la somme des pixels dans la région la plus claire. Ceci peut être accompli en utilisant des fonctionnalités similaires à Haar [34].

Une caractéristique de type Haar est représentée en prenant une partie rectangulaire d'une image et en divisant ce rectangle en plusieurs parties. Ils sont souvent visualisés sous forme de rectangles adjacents en noir et blanc [34] :

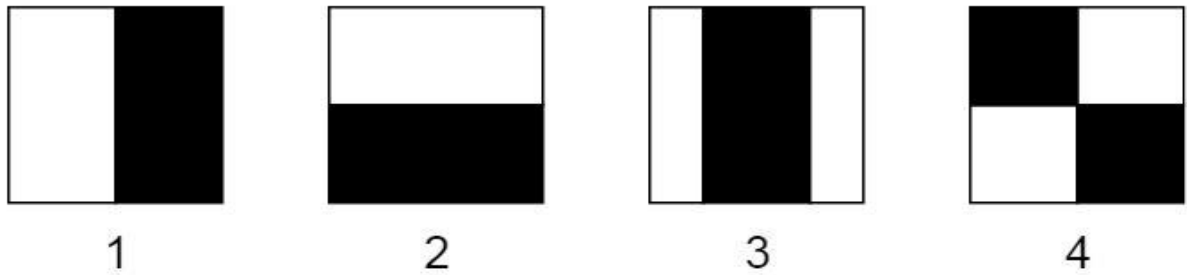


Figure III.2. Caractéristiques de base de type rectangle Haar [34].

Une caractéristique est calculée comme suit : la somme des valeurs de pixels dans la zone noire moins la somme des valeurs de pixels dans la zone blanche. Pour les zones uniformes comme un mur, ce nombre serait proche de zéro et ne vous donnerait aucune information utile [34].

Les caractéristiques sont calculées à toutes les positions et à toutes les échelles dans une fenêtre de détection de petite taille, typiquement de 24x24 pixels ou de 20x15 pixels [8].

Un très grand nombre de caractéristiques par fenêtre est ainsi généré, Viola et Jones donnant l'exemple d'une fenêtre de taille 24 x 24 qui génère environ 160 000 caractéristiques [8].

Cela devient rapidement coûteux en calcul puisque ça prend beaucoup de temps en utilisant les ressources limitées d'un ordinateur. Pour résoudre ce problème, Viola et Jones ont utilisé des images intégrales [34].

➤ L'image intégrale

Les auteurs ont proposé une nouvelle méthode qui permet de détecter rapidement les objets : l'image intégrale [33].

C'est une représentation sous la forme d'une image, de même taille que l'image d'origine, elle contient en chacun de ses points la somme des pixels situés au-dessus et à gauche du pixel courant. Plus formellement, l'image intégrale ii au point (x, y) est définie à partir de l'image i par [8]:

$$ii(x, y) = \sum_{x' < x, y' < y} i(x', y')$$

Le calcul de la somme des valeurs des pixels appartenant à une zone rectangulaire s'effectue donc en accédant seulement à quatre pixels de l'image intégrale : soit un rectangle ABCD dont les sommets sont nommés dans le sens des aiguilles d'une montre en commençant par le sommet supérieur gauche et soit x la valeur sous la représentation intégrale d'un sommet X du rectangle ($X \in \{A, B, C, D\}$). La somme des valeurs des pixels appartenant à ABCD est, quelle que soit sa taille, donnée par $C - B - D + A$ [8].

Une caractéristique de Haar étant une combinaison linéaire de tels rectangles ABCD, son calcul se fait alors en un temps indépendant de sa taille [8].

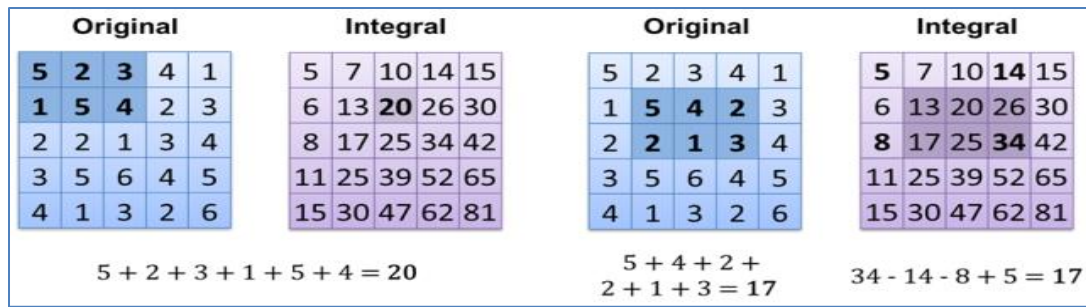


Figure III.3. La représentation d'une image intégrale [8].

➤ **Algorithme d'apprentissage basé sur Adaboost**

Le deuxième élément clé de la méthode de Viola et Jones est l'utilisation d'une méthode de boosting afin de sélectionner les meilleures caractéristiques. Le boosting est un principe qui consiste à construire un classifieur fort à partir d'une combinaison pondérée de classifieurs faibles, c'est-à-dire, donnant en moyenne une réponse meilleure qu'un tirage aléatoire. Viola et Jones adaptent ce principe en assimilant une caractéristique à un classifieur faible, en construisant un classifieur faible qui n'utilise qu'une seule caractéristique [8].

L'apprentissage du classifieur faible consiste alors à trouver la valeur seuil de la caractéristique qui permet de mieux séparer les exemples positifs des exemples négatifs. Le classifieur se réduit alors à un couple (caractéristique, seuil) [8].

L'algorithme de boosting utilisé est en pratique une version modifiée d'AdaBoost, qui est utilisée à la fois pour la sélection et pour l'apprentissage d'un classifieur fort. Les classifieurs faibles utilisés sont souvent des arbres de décision. Un cas remarquable, fréquemment rencontré, est celui de l'arbre de profondeur 1, qui réduit l'opération de classification à un simple seuillage [8].

L'algorithme est de type itératif, à nombre d'itérations déterminé. À chaque itération, l'algorithme sélectionne une caractéristique, qui sera ajoutée à la liste des caractéristiques sélectionnées aux itérations précédentes, et le tout va contribuer à la construction du classifieur fort final. Cette sélection se fait en entraînant un classifieur faible pour toutes les caractéristiques et en sélectionnant celui avec l'erreur la plus faible sur l'ensemble d'apprentissage [8].

L'algorithme tient également à jour une distribution de probabilité sur l'ensemble d'apprentissage, réévaluée à chaque itération en fonction des résultats de classification [8].

En particulier, plus de poids est attribué aux exemples difficiles à classer, c'est-à-dire ceux dont l'erreur est élevée. Le classifieur fort final construit par AdaBoost est composé de la somme pondérée des classifieurs sélectionnés [8].

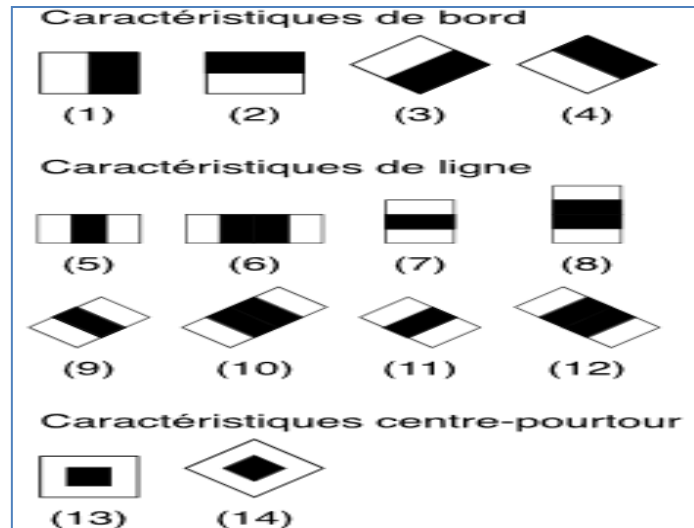


Figure III.4. Caractéristiques pseudo-haar avec différentes orientations [8].

➤ **Cascade de classifieurs**

Cascade classifieur se focalise sur les objets comme le montre la Figure III.2. précédente C'est un mécanisme qui reconnaît les régions voulues et rejette les surfaces qui n'incluent aucun objet ceci est très rapide en réalité en temps réel de détection [34].

La méthode de Viola et Jones est basée sur une approche par recherche exhaustive sur l'ensemble de l'image, qui teste la présence de l'objet dans une fenêtre à toutes les positions et à plusieurs échelles. Cette approche est cependant extrêmement coûteuse en calcul. L'une des idées clés de la méthode pour réduire ce coût réside dans l'organisation de l'algorithme de détection en une cascade de classifieurs. Appliqués séquentiellement, ces classifieurs prennent une décision d'acceptation si la fenêtre contient l'objet, l'exemple est alors passé au classifieur suivant, ou de rejet si la fenêtre ne contient pas l'objet et dans ce cas l'exemple est définitivement écarté [8].

L'idée est que l'immense majorité des fenêtres testées étant négatives (c.-à-d. ne contiennent pas l'objet), il est avantageux de pouvoir les rejeter avec le moins possible de calculs. Ici, les classifieurs les plus simples, donc les plus rapides, sont situés au début de la cascade, et rejettent très rapidement la grande majorité des exemples négatifs. Cette structure en cascade peut également s'interpréter comme un arbre de décision dégénéré, puisque chaque nœud ne comporte qu'une seule branche [8].

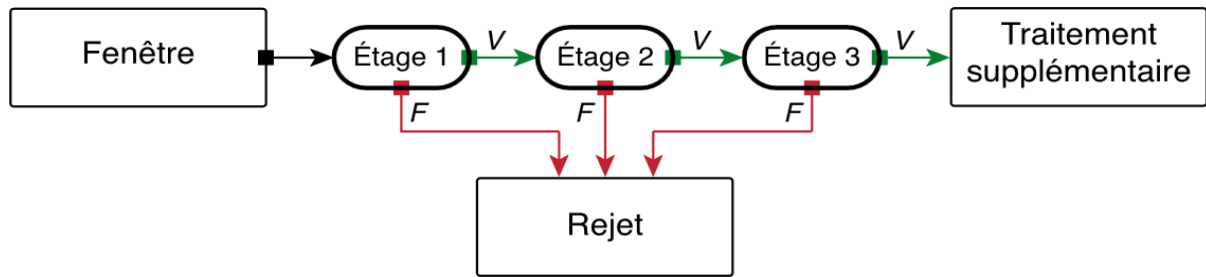


Figure III.5. Architecture de la cascade [8].

III.3. Extraction des points caractéristiques 2D et 3D

➤ Par la méthode du maillage 3D

- **Alignement 2D**

L'alignement 2D commence son processus d'alignement en détectant quelques points de repères à l'intérieur de la culture de détection, centrés au centre des yeux, de la pointe du nez et de la bouche [35].

Cette méthode d'alignement est similaire à celle utilisée dans LFW-a, qui a été fréquemment utilisée pour améliorer la précision de la reconnaissance.

Ils utilisent les repères détectés pour normaliser les images de visage (via l'échelle, la rotation et la traduction) [35].

- **Alignement des visages dense basé sur la régression en cascade**

Les travaux antérieurs sur l'alignement des faces basé sur la régression en cascade se limitaient à un petit nombre de repères de référence. Nous obtenons un alignement dense avec une taille de modèle gérable [36].

Nous montrons que cela est réalisable en utilisant un nombre relativement petit de mesures éparées et une représentation comprimée des mises à jour de déplacement de points de repère [36].

En outre, les repères faciaux sont toujours cohérents dans toutes les positions, ce qui élimine les divergences entre les annotations 2D et 3D qui ont nui aux approches précédentes [36].

➤ **Alignement 3D**

Généralement le modèle du visage est représenté par des images 2.5D et 3D (voir Figure III.6), puisque l'alignement 2D ne permet de normaliser les variations que dans le plan 2D (par exemple, voir une face depuis son côté gauche / droit). Pour normaliser les variations hors plan, il faut passer par cette transformation 3D [35].

Dans ce cas, plusieurs points de repère supplémentaires sont détectés sur les faces. Puis, chaque pixel dans le plan X-Y stocke la valeur de sa profondeur z, ce qui nous donne une représentation bidimensionnelle d'un ensemble de points 3D (x,y,z) qui correspond à une image 2.5D (image de profondeur) [19].

[Chapitre III : Les Méthodes Utilisées]

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 [19]. Après avoir connecté cette liste de points par des arêtes (polygones), nous aurons une représentation surfacique par lot de facettes ou maillage de polygones. En effet c'est un ensemble de facettes polygonales, généralement triangles ou quadrangles qui ne sont pas nécessairement structurés [32], et c'est la méthode la plus simple pour représenter un visage 3D. [19].

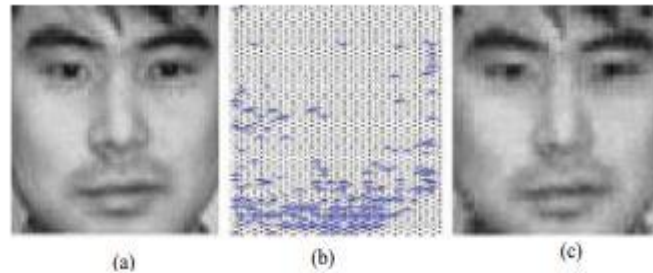


Figure III.6. Exemple : (a) une image du visage originale, (b) sa projection (SOM-Face) et (c) image reconstruite [19].

Mis à part notre technique, il existe d'autres techniques pour construire un maillage 3D comme les techniques de combinaison des images 2.5D ou celles qui exploitent des systèmes d'acquisition 3D tel que le scanner 3D comme le montre la figure III.7 [19].

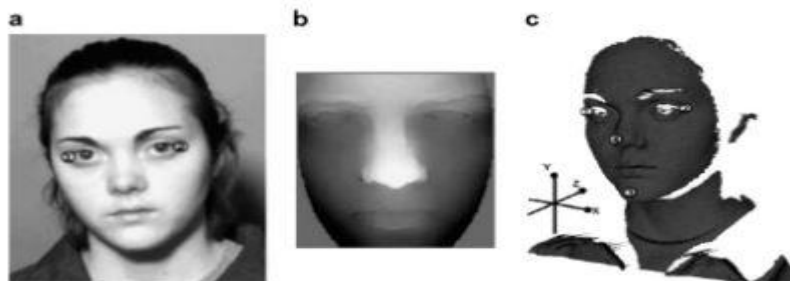


Figure III.7. (a) Image texture (b) Image 2.5D (c) Image 3D [19].

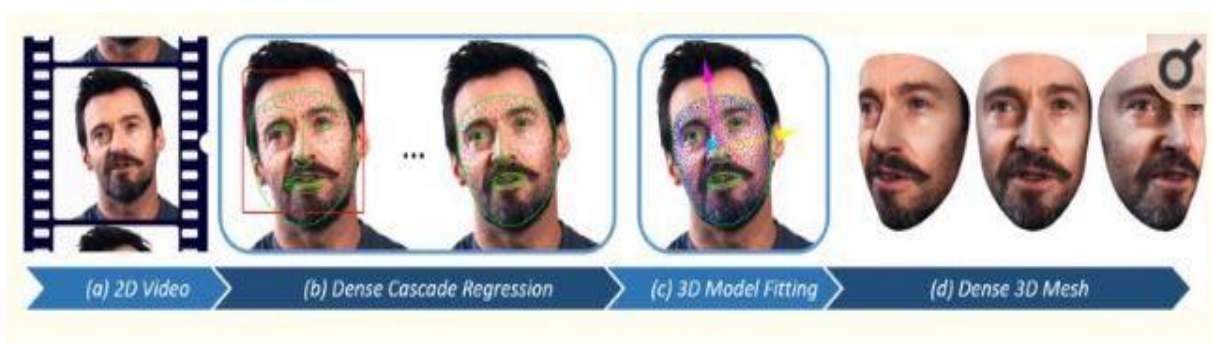


Figure III.8. De la 2D vers la 3D [36].

[Chapitre III : Les Méthodes Utilisées]

Dans la figure III.8. En partant d'une image 2D du visage d'une personne (a), un ensemble dense de repères faciaux est estimé à l'aide d'un cadre de régression en cascade rapide et cohérent (b), puis un modèle 3D déformable basé sur les parties est appliqué (c) pour reconstruire un maillage 3D dense du visage (d) [36].

III.4. Classification des visages 3D

➤ Par la méthode de CNN

En 2012, une révolution se produit : lors de la compétition annuelle de vision par ordinateur ILSVRC, un nouvel algorithme de Deep Learning explose les records ! Il s'agit d'un **réseau de neurones convolutif** [37].

TensorFlow est un framework créé par Google pour la création de modèles Deep Learning. Deep Learning est une catégorie de modèles d'apprentissage automatique (= algorithmes) utilisant des réseaux de neurones multicouches [44].

Les réseaux de neurones convolutifs ont une méthodologie similaire à celle des méthodes traditionnelles d'apprentissage supervisé : ils reçoivent des images en entrée, détectent les *features* de chacune d'entre elles, puis entraînent un classifieur dessus [37].

Les réseaux de neurones convolutionnels proposés initialement par Le Cun. Ce choix a été motivé principalement par ce qu'il intègre implicitement une phase d'extraction de caractéristiques et il a été utilisé avec succès dans de nombreuses applications [38].

Ils sont réputés pour leur robustesse aux faibles variations d'entrée, le faible taux de prétraitement nécessaires à leur fonctionnement [38].

Le CNN est un réseau de neurone multicouche qui est spécialisés dans des tâches de reconnaissance de forme. Ces réseaux ont été inspirés par les travaux de Hubel et Wiesel sur le cortex visuel chez les mammifères qui combine trois idées principales [38] :

- Les champs récepteurs locaux.
- Les poids partagés.
- Le sous-échantillonnage.

L'architecture de CNN repose sur plusieurs réseaux de neurones profonds consistant en une succession de couches de convolution et d'agrégation (pooling) est dédié à l'extraction automatique de caractéristiques, tandis que la seconde partie, composée de couches de neurones complètement connectés, est dédiée à la classification [38].

Chaque cellule des couches de convolution est connectée à un ensemble de cellules regroupées dans un voisinage rectangulaire sur la couche précédente. Les champs récepteurs locaux permettent d'extraire des caractéristiques basiques [38].

Les couches sont dites « à convolution » car les poids sont partagés et chaque cellule de la couche réalise la même combinaison linéaire (avant d'appliquer la fonction sigmoïde) qui

[Chapitre III : Les Méthodes Utilisées]

peut être vue comme une simple convolution. Ces caractéristiques sont alors combinées à la couche suivante afin de détecter des caractéristiques de plus haut niveau [38].

Entre deux phases d'extraction de caractéristiques, le réseau réduit la résolution de la carte des caractéristiques par un moyen de sous-échantillonnage. Cette réduction se justifie à deux titres : diminuer la taille de la couche et apporter de la robustesse par rapport aux faibles distorsions [38].

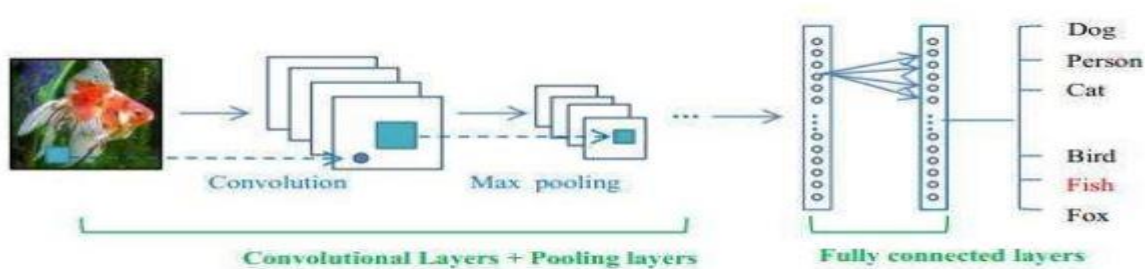


Figure III.9. L'architecture d'un réseau de neurone convolutif [38].

III.4.1. L'apprentissage en profondeur

L'apprentissage en profondeur est une technique permettant de mettre en œuvre l'apprentissage automatique. Il utilise parfois des réseaux de neurones pour apprendre. L'utilisation d'arbres de décision peut également être qualifiée d'apprentissage en profondeur, mais l'apprentissage en profondeur implique généralement l'utilisation de réseaux de neurones [37].

III.4.2. Couche de convolution

La convolution est une opération mathématique comme l'addition et la multiplication, il est très utile de simplifier des équations plus complexes, cette opération est largement utilisée dans le traitement du signal numérique.

Lorsqu'on applique la convolution aux traitements d'image, on combine l'image d'entrée avec une sous-région de cette image (filtre). Le filtre est aussi connu sous le nom du noyau de convolution, il consiste en des poids de cette sous-région. La sortie de cette couche est l'image entrée avec des modifications qui est souvent appelée une carte de caractéristique (feature Map) [38].

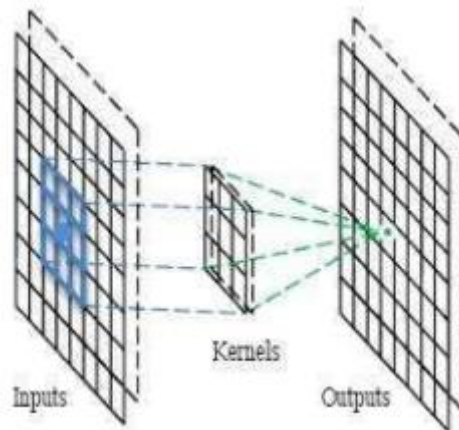


Figure III.10. L'opération de convolution [38].

En terme mathématique, Une couche de convolution C_i (couche i du réseau) est paramétrée par son nombre N de cartes de convolution $M_{j i}$ ($j \in \{1, \dots, N\}$), la taille des noyaux de convolution $K_x \times K_y$ (souvent carrée), et le schéma de connexion à la couche précédente L_{i-1} [38].

Chaque carte de convolution est le résultat d'une somme de convolution des cartes de la couche précédente par son noyau de convolution respectif.

Un biais est ensuite ajouté et le résultat est passé à une fonction de transfert non-linéaire. Dans le cas d'une carte complètement connectée aux cartes de la couche précédente, le résultat est alors calculé par [38] :

$$[31] = + \sum [31] \quad 3.2.2$$

III.4.3. Couche de sous-échantillonnage (Pooling)

Dans les architectures classiques de réseaux de neurones convolutionnels, les couches de convolution sont suivies par des couches de sous échantillonnage (couche d'agrégation). Cette dernière réduit la taille des cartes de caractéristique pour but de diminuer la taille de paramètre, et renvoie les valeurs maximales des régions rectangulaires de son entrée [37].

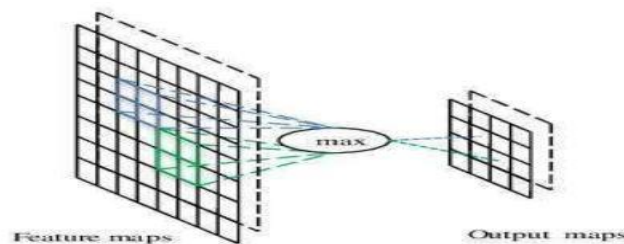


Figure III.11. L'opération de sous échantillonnage [38].

[Chapitre III : Les Méthodes Utilisées]

Couche entièrement connectée : Les paramètres des couches de convolution et de max agrégation sont choisis de sorte que les cartes d'activation de la dernière couche soient de taille 1, ce qui résulte en un vecteur 1D d'attributs. Des couches classiques complètement connectées composées de neurones sont alors ajoutées au réseau pour réaliser la classification.

La dernière couche, dans le cas d'un apprentissage supervisé, contient autant de neurones que de classes désirées. Cette couche contient N neurones, et une fonction d'activation de type sigmoïde est utilisée afin d'obtenir des probabilités d'appartenance à chaque classe [37].

III.5 Conclusion

Dans ce chapitre nous avons cités les méthodes les plus connues de la détection de visage et reconnaissance du visage 3D que nous allons adopter pour implémenter notre application, en ce qui concerne la détection d'un visage nous avons choisi la méthode de Viola-Jones pour sa performance et le maillage 3D pour sa simplicité pour représenter un visage 3D, les **réseaux de neurones convolutifs** pour leur robustesse et leur précision. Dans la partie qui suit nous allons présenter la conception et la réalisation de notre projet.

Chapitre IV

La Conception et La Réalisation

[Chapitre IV : La Conception et La Réalisation]

IV.1. Introduction

Après avoir donné les notions de base sur de la reconnaissance faciale en 3D, il s'agit maintenant de mettre en œuvre les étapes explicitées dans les chapitres précédents pour obtenir une application de « Reconnaissance faciale en 3D ». Ce chapitre porte sur la construction d'une application qui permet de reconnaître individu en utilisant ls réseaux de neurones convolutionnels.

IV.2. Environnement matériel et logiciel

IV.2.1. Ressources utilisées

Les ressources physiques utilisées sont :

- Processeur Intel® Core™ i3-5005U d'une fréquence de 2.0 GHz.
- Une mémoire vive d'une capacité de 4 GO.
- Une carte graphique Intel® HD 5500M.

Et pour ce qui est côté logiciel (Soft) :

- Système d'exploitation : Linux Ubuntu 16.04
- Langage de programmation : Python3.6.

IV.2.2. Langage de programmation

Le langage de programmation choisi dans notre travail est Python qui reste un des langages les plus utilisés dans le domaine de la reconnaissance des formes. Cela est dû au fait que le langage Python est facile à apprendre, à lire, à comprendre et à écrire. Le langage Python est portable (fonctionne sous de nombreux systèmes d'exploitation) ; doté d'une communauté active [39].

Python est un langage de programmation puissant et facile à apprendre. Il dispose de structures de données de haut niveau et permet une approche simple mais efficace de la programmation orientée objet. Parce que sa syntaxe est élégante, que son typage est dynamique et qu'il est interprété, Python est un langage idéal pour l'écriture de scripts et le développement rapide d'applications dans de nombreux domaines et sur la plupart des plateformes [39].

L'interpréteur Python et sa vaste bibliothèque standard sont disponibles librement, sous forme de sources ou de binaires, pour toutes les plateformes majeures depuis le site Internet <https://www.python.org/> et peuvent être librement redistribués. Ce même site distribue et pointe vers des modules, des programmes et des outils tiers. Enfin, il constitue une source de documentation [39].

L'interpréteur Python peut être facilement étendu par de nouvelles fonctions et types de données implémentés en C ou C++ (ou tout autre langage appellable depuis le C). Python est également adapté comme langage d'extension pour personnaliser des applications [39].

[Chapitre IV : La Conception et La Réalisation]

➤ PyQT4

PyQt est un module Python qui permet d'accéder à la bibliothèque Qt de Nokia. Il s'agit d'un framework applicatif multiplateforme : une même application PyQt fonctionne sans problème sous OS X, Linux et Windows.

C'est aussi un gros module de plus de trois cent classes, de six mille fonctions. Il fournit un ensemble complet de classes utiles qui fournissent des interfaces pour l'accès aux fichiers, parser du XML, des structures de données et plus, ainsi que des outils comme Qt Designer, Qt Assistant et Qt Linguist pour concevoir des interfaces, accéder à l'aider et faciliter les traductions.

C'est donc un framework complet : il fournit des classes qui gèrent les types basiques de données comme celles qui effectuent des tests unitaires. PyQt est développé par Phil Thompson, chez Riverbank Computing Limited [40].

Les composants majeurs de PyQt sont :

- Le module QtGui, pour tout ce qui concerne les GUI.
- Le module QtCore, pour tout ce qui ne concerne pas les GUI.
- L'espace de noms Qt, pour les types et autres identifiants.
- Les outils Qt, pour d'autres tâches variées.

D'autres modules plus spécialisés sont aussi disponibles : QtHelp, QtOpenGL, QtScript, QtXml, QSql, QTest et bien d'autres. Des scripts pour les ressources, interfaces et traductions sont aussi fournis, pour compatibilité avec son pendant C++ [40].

La meilleure manière d'apprendre à utiliser Qt est de suivre la documentation, par exemple avec Qt Assistant, qui vient avec lui. Bien que prévue pour le C++, l'utilisation que l'on en fait en Python n'est pas très différente. Qt n'est pas seulement partout sur toutes les plateformes, il est aussi sur tous les langages, Ruby, Python, C++, Java et d'autres [40].

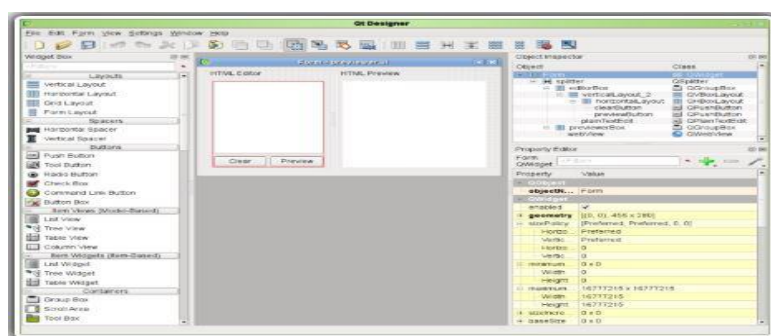


Figure IV.1. Interface de PyQt4 [40].

IV.3. Système « FaceIt »

Pour nommer notre système nous nous sommes inspirés du fait que l'individu se mette en face de la Webcam pour détecter son visage (y faire face, « face it » en anglais).

[Chapitre IV : La Conception et La Réalisation]

FaceIt effectuera la reconnaissance du visage et affichera le résultat obtenu. Notre système possède les caractéristiques suivantes :

- Il détecte le visage de l'individu.
- Il extrait les points caractéristiques de ce dernier.
- Il est sensible aux variations des expressions faciales.

IV.3.1. Méthodes de reconnaissance faciale 3D

Le tableau ci-dessus présente le taux de reconnaissances faciales obtenus par plusieurs applications qui utilisent des méthodes différentes :

Sources	Années	Applications	Méthodes utilisées	Taux de reconnaissance %
Lahoucine Ballihi, Université des Sciences et Technologie, Lille [26]	2012	Application à la reconnaissance des visages et à la classification du genre	CNN réseaux de neurone convolutionnels	86,05
Khefif Bouchra Université Abou Bakr Belkaid – Tlemcen Faculté des Sciences [30]	2013	Mise au point d'une application de reconnaissance faciale	ACP analyse en composantes principales	60
BEGGARI Samiha et KHAMRA Khaoula Université Kasdi Merbah, Ouargla [38]	2017	Système de reconnaissance de visage par un réseau de neurone convolutionnel	CNN réseaux de neurone convolutionnels	80
Wassila Sghier, Encadrante : C.M Bentaouza, Université Abdelhamid Ibn Badis, Mostaganem [8]	2016	Reconnaissance d'Expressions Faciale 3D	Approche 3D dynamique	93.83
P. Lemaire1, P. Szeptycki1, M. Ardabilian1, L. Chen Ecole Centrale de Lyon, Département Math-Info, France [41]	2014	Reconnaissance de visages en 3D orientée région	ICP	78.5%

Tableau IV. Taux de reconnaissance.

IV.4. Organigramme du système de reconnaissance du visage en 3D

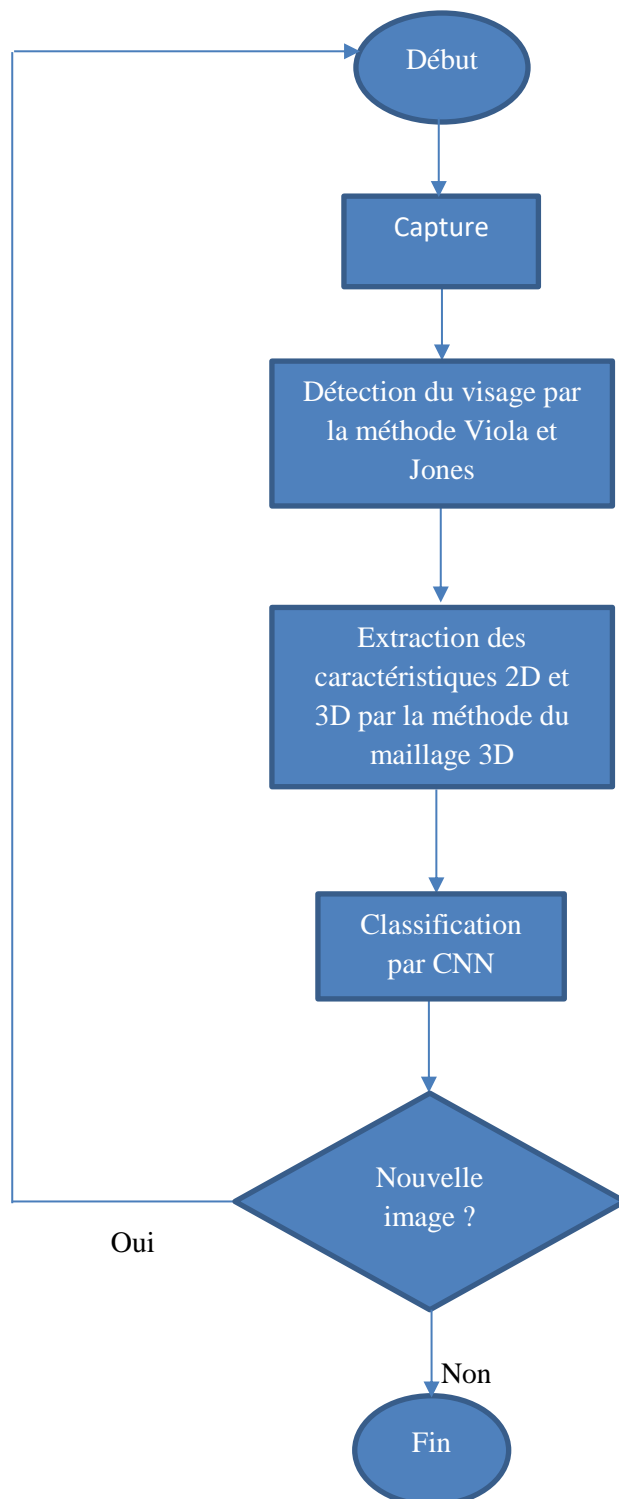


Figure IV.2. Organigramme de système de reconnaissance du visage en 3D.

IV.5. Présentation de l'application

Nous présentons dans cette section les différents aspects du système d'identification faciale en 3D.

IV.5.1. Fonctionnement du système

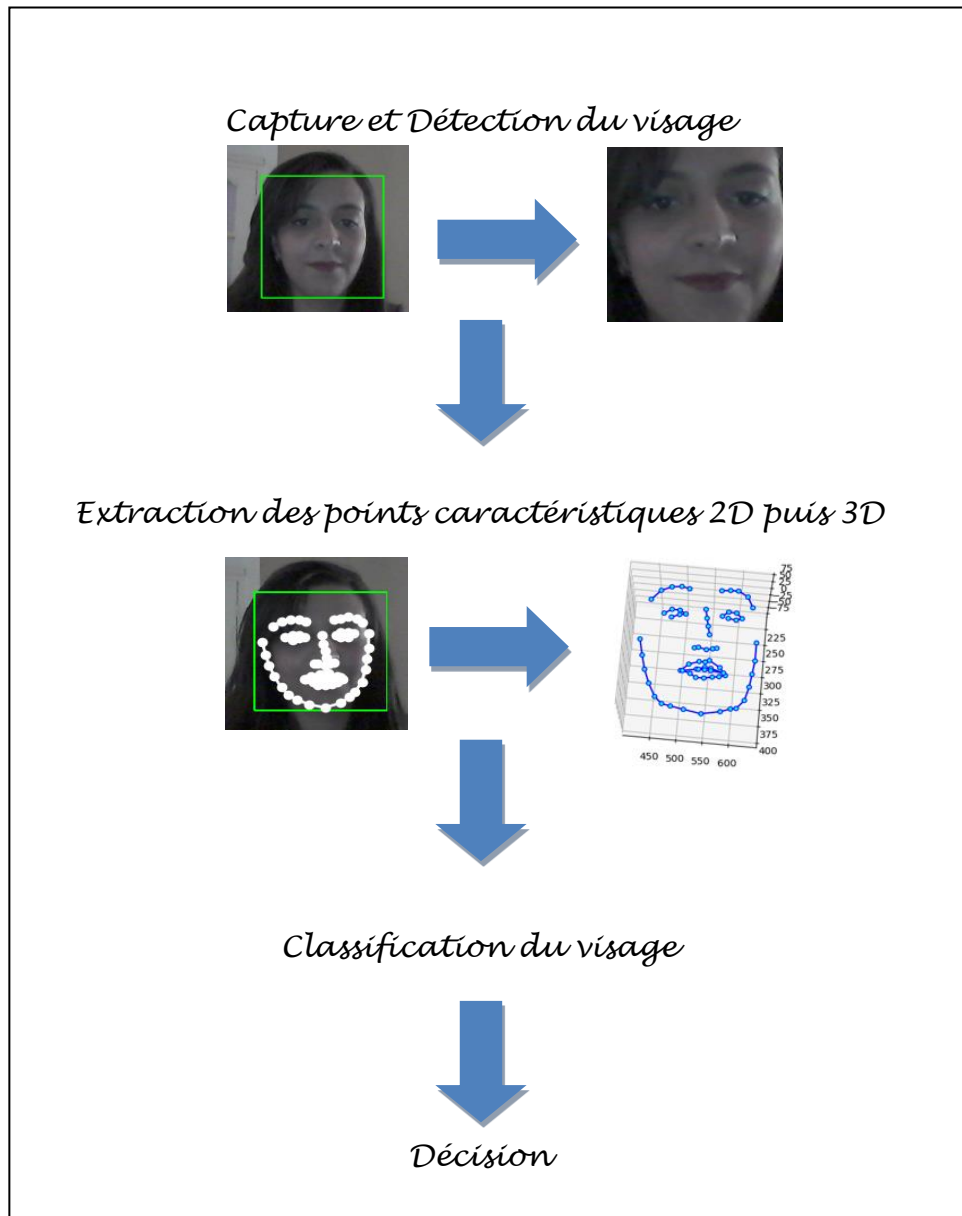


Figure IV.3. Schéma du système « *FaceIt* »

IV.5.2. Capture

La capture des images sont fait à l'aide d'une caméra Webcam intégrée à l'ordinateur de marque Acer :



Figure IV.4. Caméra utilisée.

IV.5.3. Interface principale

Comme vous pouvez le voir sur la figure qui suit est l'interface initiale de notre application.

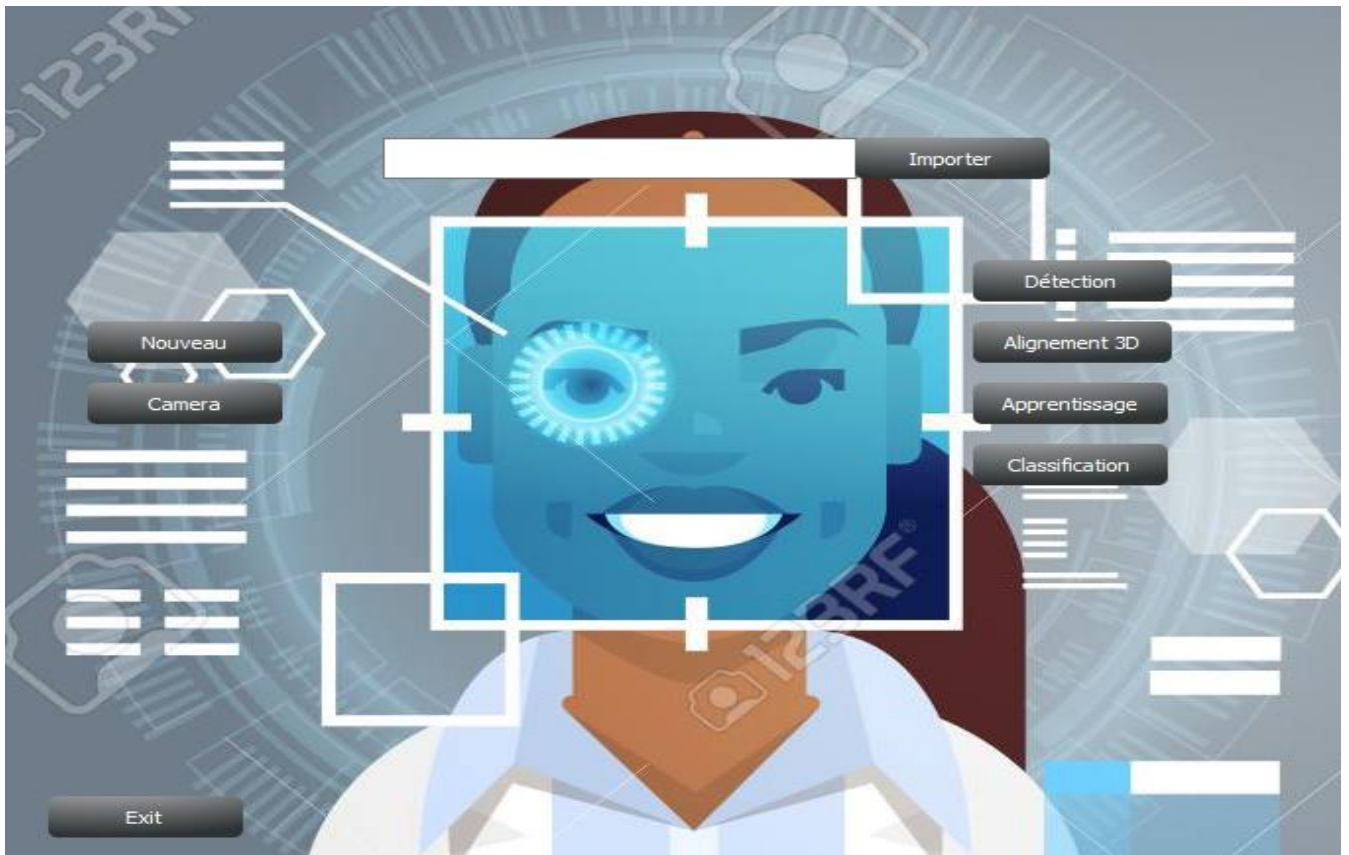


Figure IV.5. Interface principale.

[Chapitre IV : La Conception et La Réalisation]

IV.5.4. Capture

En utilisant la bibliothèque « OpenCv » on a réussie à détecter un visage en dessinant dessus un petit rectangle vert.

En cliquant sur le bouton Caméra, la caméra s'allume, cela nous permet de détecter un visage comme le montre la figure suivante :

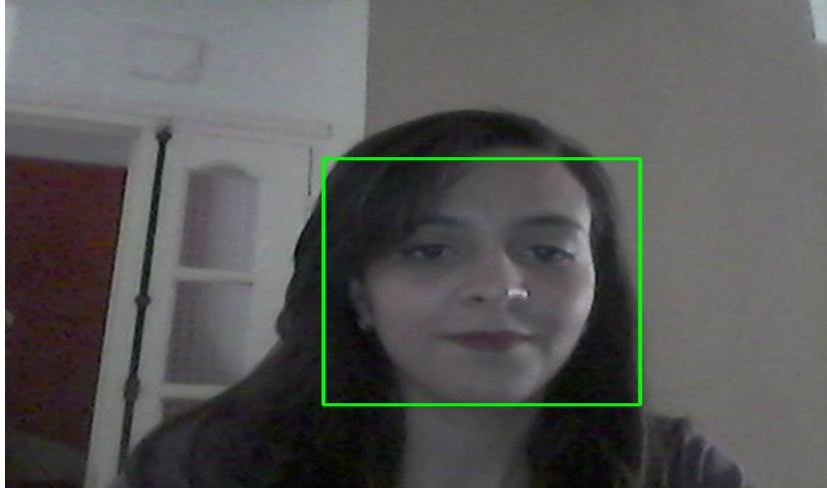


Figure IV.6. Détection du visage.

IV.5.5. Détection du visage

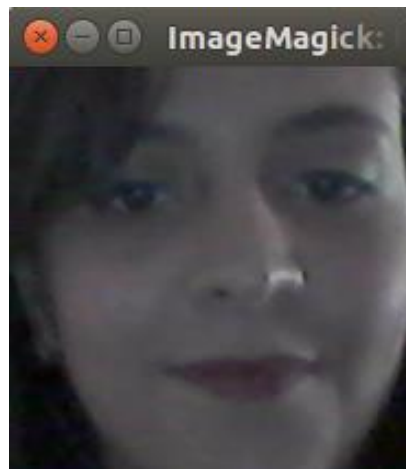


Figure IV.7. Visage détecté.

IV.5.6. Extraction des caractéristiques 2D et 3D

En cliquant sur le bouton Alignment 3D, cela nous permet de transformer le visage en 3D ainsi qu'extraire les caractéristiques faciales de ce dernier.

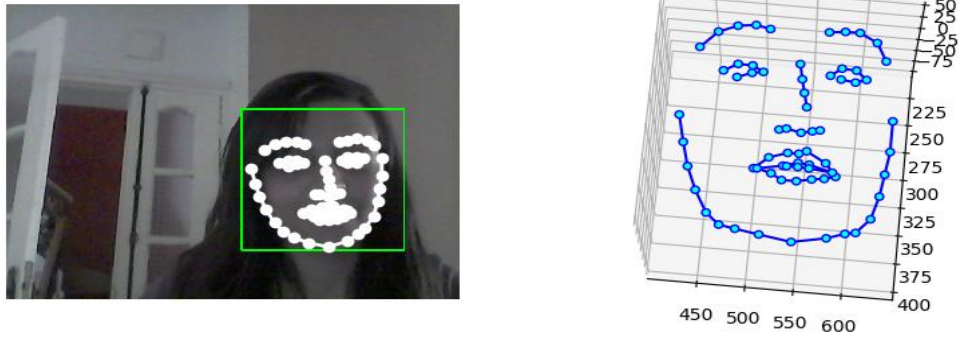


Figure IV.8. Extraction des caractéristiques faciales.

IV.5.7. Fichier JSON

La figure ci-dessous représente un fichier JSON qui contient les données vectorielles des images sur lesquelles nous faisons notre traitement.

Lorsque nous fournissons une image d'entrée au modèle, cela nous donne 128 octets de données vectorielles numériques à l'image.

```
1 {"laroui": {"Left": [[0.17641136050224304, -0.07008262723684311, -0.1818150132894516, 0.08612490445375443, 0.07453210651874542,
-0.10545285046100616, 0.074374258518219, -0.12702126801013947, 0.03300271928310394, 0.07762300223112106, -0.060170531272888184,
-0.006257229950278997, 0.03748017176985741, -0.17277821898460388, 0.10500302910804749, 0.06462818384170532, 0.13625065982341766,
0.1095799058675766, 0.08422254025936127, -0.01277566235512495, -0.05838850885629654, 0.027825644239783287, -0.02588777057826519,
0.11119961738586426, 0.05948071554303169, 0.10222112387418747, -0.012659572996199131, -0.020331352949142456, 0.019338499754667282,
-0.022403256967663765, 0.1524747759103775, 0.07054413110017776, -0.024055326357483864, 0.02073703147470951, 0.11312606930732727,
-0.10663534700870514, 0.014348329044878483, 0.07047254592180252, 0.07996346056461334, -0.06832391768693924, -0.11466734111309052,
0.1392854005098343, 0.08849440515041351, -0.07458731532096863, -0.07376844435930252, 0.023711461573839188, 0.047428373247385025,
-0.033956628292798996, -0.0919555127620697, 0.07698666304349899, 0.030016576871275902, -0.0830840989947319, -0.08821078389883041,
0.10407537966966629, 0.16530652344226837, -0.032069507986307144, -0.0399915836751461, 0.04931722953915596, -0.039074335247278214,
-0.13593553006649017, -0.0015465214382857084, -0.04520250856876373, 0.07183056324720383, 0.002822855021804571, 0.030692633241415024,
0.12735410034656525, -0.019441472366452217, 0.06181243434548378, 0.01237778551876545, 0.07719828933477402, 0.13418996334075928,
-0.1521591693162918, 0.028850531205534935, -0.0039069764316082, -0.1042969822883606, 0.01028698030859232, -0.09721103310585022,
-0.06947754323482513, 0.005463581532239914, -0.1289377212524414, 0.07977964729070663, -0.06467583775520325, -0.03965795412659645,
0.1348368376493454, -0.017166074365377426, -0.08874253183603287, 0.051489073783159256, -0.1331547498703003, -0.049231547862291336,
-0.023254981264472008, 0.05245346948504448, -0.05054326355457306, -0.11667054891586304, -0.0413406603038311, -0.08947151154279709,
-0.02694661356508732, 0.0032652136869728565, 0.02729189209640026, -0.1271933913230896, -0.06200943887233734, -0.11454205960035324,
0.145210862159729, -0.012304866686463356, -0.12735961377620697, 0.03073952533046913, 0.031080646440386772, -0.06946928054094315,
-0.08032447099685669, -0.032746441662311554, -0.028194021433591843, -0.09942030161619186, 0.1424144208431244, 0.06076119840145111,
-0.11946091800928116, -0.009942350909113884, -0.008629062213003635, -0.23000971972942352, -0.0856660008430481, 0.050670091062784195,
-0.022036194801330566, -0.05003717169165611, -0.026781149208545685, 0.15608343482017517, 0.06160449981689453, 0.13730275630950928,
-0.09751643240451813, 0.05810728669166565, -0.11805766075849533]], "Right": [[0.15704257786273956, -0.07130566984415054,
-0.11148577183485031, 0.033673014491796494, 0.12666936218738556, -0.12484721094369888, 0.042941126972436905, -0.1147739514708519,
-0.03385376185178757, 0.08337804675102234, -0.021568838506937027, -0.04389769583940506, -0.015368398278951645, -0.14792147278785706,
0.08829596638679504, 0.05091138184070587, 0.11608816683292389, 0.09007011353969574, 0.03580385819077492, -0.022042931988835335,
-0.0877063199877739, 0.04879974573850632, 0.03685925528407097, 0.11482647061347961, 0.11395785212516785, 0.14921489357948303,
0.030134163796901703, 0.04412205517292023, -0.004568207077682018, -0.027868669480085373, 0.13861502707004547, 0.06811460852622986,
-0.0679260641336441, 0.025521468371152878, 0.12786883115768433, -0.04746483266353607, -0.0018703198293223977, 0.04905445873737335,
0.08748427033424377, -0.015790412202477455, -0.1036970466375351, 0.08000358939170837, 0.10830611735582352, -0.040235474705696106,
-0.043299350887537, -0.003837407799437642, 0.03752390295267105, -0.08092769235372543, -0.10489988327026367, 0.0689566433429718,
-0.0399014912545681, -0.034153349697589874, -0.014064235612750053, 0.06109645590186119, 0.07740572839975357, 0.015597349964082241,
0.02437092736363411, 0.03821374848484993, -0.057839009910821915, -0.1685638576745987, 0.04308966174721718, -0.06318619102239609,
0.08691339939832687, 0.0823601558804512, -0.0334443524479866, 0.10119105130434036, -0.007479275576770306, 0.029049992561340332,
0.0525832399725914, 0.08288940042257309, 0.08684150129556656, -0.1441170573234558, 0.055176638066768646, -0.025025993585586548,
-0.06650514900684357, -0.018161287531256676, -0.06598129868507385, -0.03752923011779785, -0.0033139700535684824, -0.1233893483877182,
0.0941271260380745, -0.0476694293320179, -0.039085131138563156, 0.19451889395713806, 0.037264786660671234, -0.07719854265451431,
0.0863676369190216, -0.14625690877437592, -0.04479997605085373, 0.02582259103655815, 0.01718844659626484, -0.08541284501552582,
-0.12739238142967224, -0.04593212902545929, -0.026711929589509964, -0.04309780150651932, -0.009762720204889774, -0.04165080934762955,
```

Figure IV.9. Fichier JSON.

IV.5.8. Apprentissage

En cliquant sur le bouton Apprentissage, le script Train.sh est exécuter ainsi l'apprentissage est lancé.

Après l'exécution de la phase d'apprentissage, nous avons obtenu un taux de précision de 81.8%.

```
jojo@jojo-Aspire-ES1-571: ~/Bureau/Appli
Instructions for updating:
Future major versions of TensorFlow will allow gradients to flow
into the labels input on backprop by default.
See `tf.nn.softmax_cross_entropy_with_logits_v2`.
2019-06-03 17:36:49.868024: Step 0: Train accuracy = 42.0%
2019-06-03 17:36:49.913833: Step 0: Cross entropy = 2.002324
2019-06-03 17:36:51.183262: Step 0: Validation accuracy = 29.0% (N=100)
2019-06-03 17:36:52.612476: Step 10: Train accuracy = 64.0%
2019-06-03 17:36:52.612565: Step 10: Cross entropy = 1.616519
2019-06-03 17:36:52.750822: Step 10: Validation accuracy = 48.0% (N=100)
2019-06-03 17:36:53.821215: Step 20: Train accuracy = 81.0%
2019-06-03 17:36:53.821300: Step 20: Cross entropy = 1.409156
2019-06-03 17:36:53.968820: Step 20: Validation accuracy = 78.0% (N=100)
2019-06-03 17:36:54.985979: Step 30: Train accuracy = 89.0%
2019-06-03 17:36:54.986066: Step 30: Cross entropy = 1.142486
2019-06-03 17:36:55.090219: Step 30: Validation accuracy = 79.0% (N=100)
2019-06-03 17:36:56.104964: Step 40: Train accuracy = 90.0%
2019-06-03 17:36:56.105049: Step 40: Cross entropy = 1.025776
2019-06-03 17:36:56.207857: Step 40: Validation accuracy = 82.0% (N=100)
2019-06-03 17:36:57.226763: Step 50: Train accuracy = 85.0%
2019-06-03 17:36:57.226849: Step 50: Cross entropy = 0.978000
```

Figure IV.10. Résultat d'apprentissage.

IV.5.9. Classification

En cliquant sur le bouton Classifier, cela nous permet de classer les individus selon les pourcentages de ressemblance.

En utilisant la bibliothèque TensorFlow, nous avons réussi à créer le modèles Deep Learning.

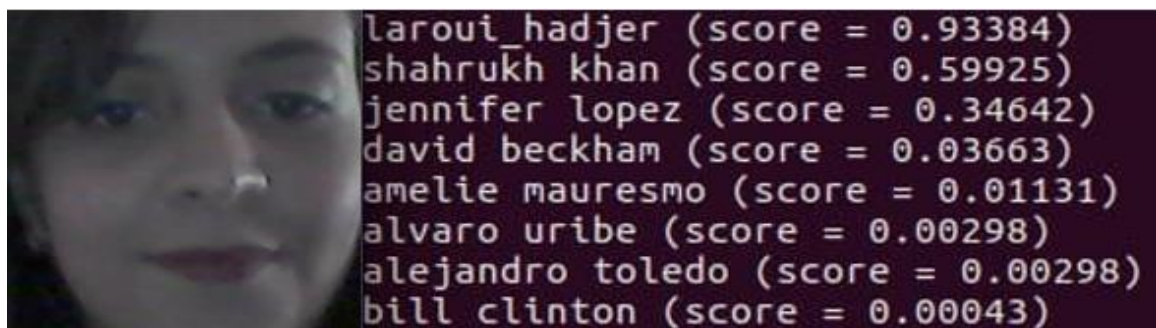
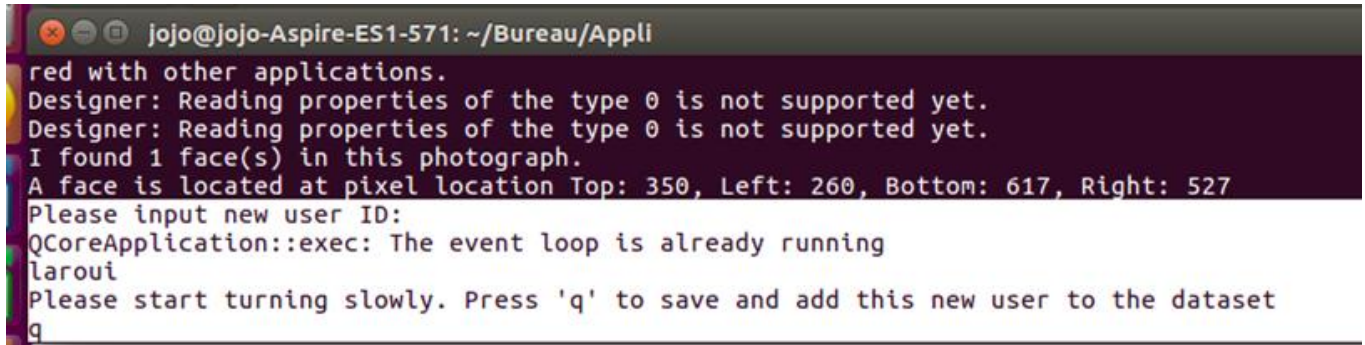


Figure IV.11. Classification de l'individu.

[Chapitre IV : La Conception et La Réalisation]

IV.5.10. Ajout d'un nouveau visage

En cliquant sur le bouton New, cela nous permet d'ajouter une nouvelle personne à notre fichier JSON.



```
jojo@Jojo-Aspire-ES1-571: ~/Bureau/Appli
red with other applications.
Designer: Reading properties of the type 0 is not supported yet.
Designer: Reading properties of the type 0 is not supported yet.
I found 1 face(s) in this photograph.
A face is located at pixel location Top: 350, Left: 260, Bottom: 617, Right: 527
Please input new user ID:
QCoreApplication::exec: The event loop is already running
laroui
Please start turning slowly. Press 'q' to save and add this new user to the dataset
q
```

Figure IV.12. Ajout d'une nouvelle personne.

IV.6. Conclusion

Dans ce chapitre, nous avons décrit les étapes de la conception du système de d'identification faciales en 3D « *FaceIt* », l'environnement matériel et logiciel, ainsi que les détails de réalisation de notre système et l'interface développée. En vue de ce qui a été fait dans la conception et la réalisation, nous pourrions dire que :

- *FaceIt* nous permet de reconnaître en quelque sorte n'importe qu'elle personne.
- *FaceIt* est performant en raison des bons taux de reconnaissances.
- Son interface est simple que toute personne ayant peu de connaissances peut l'utiliser.

Conclusion générale

Dans ce mémoire, nous nous sommes intéressés à la reconnaissance faciale, puisque cette dernière est indispensable pour assurer la sécurité dans plusieurs domaines. L'être humain a essayé de créer des applications ressemblant plus ou moins au moyen le plus puissant de détection des visages existant : la perception visuelle humaine. Cependant cette tâche reste encore complexe et difficile à réaliser.

L'objectif principal de notre projet est de concevoir et d'implémenter une application de reconnaissance faciale en 3D.

Tout d'abord, nous avons donné un bref aperçu sur la reconnaissance faciale tout en mentionnant ses trois principales étapes qui sont : la détection, l'extraction des caractéristiques et la classification. Nous avons cité les différentes approches de la détection du visage ainsi que les différents obstacles qu'a rencontrés la reconnaissance faciale.

Puis, nous nous sommes focalisées sur la reconnaissance faciale en 3D en montrant son processus. Nous avons cité et détaillé chacune de ses approches ainsi que les méthodes d'extraction d'informations et de classification les plus utilisées. Malgré tous les avantages de cette technique (déjà mentionnés) et ses performances qui ne cessent de se développer ces dernières années, la reconnaissance faciale 3D n'a pas encore pu résoudre tous les problèmes liés aux variations déjà rencontrés et atteindre la perfection tant voulue.

Par la suite, nous avons détaillé les trois méthodes que nous avons choisi d'utiliser pour notre application de reconnaissance faciale en 3D.

Viola-Jones est une méthode fiable et avancée pour faire une détection. Une fois que le visage ait été détecté, le maillage 3D reste le moyen le plus simple pour le transformer en un modèle 3D après avoir extrait les caractéristiques de ce dernier. Et enfin, le CNN est parmi les classificateurs les plus robustes pour réaliser une application performante.

Enfin, nous avons pu rassembler toutes les étapes déjà citées avec l'implémentation notre application « FaceIt », qui est un système ayant pu offrir de bons résultats.

[Bibliographie]

REFERENCES ET WEBOGRAPHIE

- [1] Mohamed Aymen FODDA, « DETECTION ET RECONNAISSANCE DE VISAGE », Ecole Supérieure Privée d'Ingénierie et de Technologies, Octobre 2010.
- [2] Serge KOMANDA BASEMA, « Identification des personnes par reconnaissance de visage pour la sécurité d'une institution bancaire », Institut supérieur pédagogique de Bukavu, 2010.
- [3] K.Bouchra, Mise au point d'une application de reconnaissance facial, université de Tlemcen, 2013.
- [4] A.J.Goldstein. L.D.Harmon and A.B.Lesk, Identification of Humman Faces, Proc.IEEE, May 1971, vol.59, No. 5, 748-760.
- [5] L.Sirovich and M.Kirby, A Low -Dimensional Procedure for the Characteri- zation of Human Faces, J. Optical Soc. Am. A, 1987, vol.4, No. 3, 559-524.
- [6] M.A.Turk and A.P.Pentland, Face Recognition using Eigenfaces, Proc. IEEE, 1991, 586-591.
- [7] Reconnaissance faciale : définition et traduction, <https://www.journaldunet.fr/business/dictionnaire-du-marketing/1208194-reconnaissance-faciale-definition-et-traduction/>, 12/12/2018.
- [8] Wassila Sghier, Encadrante : C.M Bentaouza « Reconnaissance d'Expressions Faciale 3D», Mémoire de Master en Informatique, Option : Ingénieries des Systèmes D'informations, Université Abdelhamid Ibn Badis, Mostaganem, Juin 2016.
- [9] W. Hizem et B. Dorizzi, « Capteur Intelligent pour la Reconnaissance de Visage », Thèse présentée pour l'obtention du diplôme de Docteur De L'Institut National Des Télécommunications Spécialité : Electronique/Informatique, Université Pierre et Marie Curie - Paris 6, 2009.
- [10] S.Guerfi. Authentification d'individus par reconnaissance de caractéristique biométriques liées aux visages 2D/3D, thèse doctorat, université d'Evry-val d'Essonne, France, 2008.
- [11] Cheng-Chin. et al. A novel method for detecting lips, eyes and faces in real time. Real-Time Imaging, 9(4) : 277-287, 2003.
- [12] M.Zrelli, implémentation d'une méthode de détection et suivi de visage en temps réel, Ecole royale militaire Bruxelles royaume de la Belgique, 2006/2007.
- [13] BERKANE Chahrazed et BERKANI Afef, « La détection des visages », Université Larbi Ben M'hidi, Oum El Bouaghi Juin 2014.
- [14] V.Bruce. Recognizing faces. Lawrence Erlbaum Associates, London, U.K,1988.

[Bibliographie]

- [15] A. Melakh, reconnaissance des visages en condition dégradées, thèse de doctorat, université d'Evry-val d'Essonne, France 2009.
- [16] P. J. Phillips, P. J. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone. Face recognition vendor test 2002: Evaluation report. NISTIR 6965.
- [17] H. Greenspan, J. Goldberger, I. Eshet. Mixture model for facecolor modeling and segmentation, Pattern Recognition Lett. 22 (14) 1525–1536, 2001.
- [18] Avantages – Inconvénients, <https://sites.google.com/site/3cb1sibjbs/home/a-solution-faciale-presentation-generale-1/avantages---inconvenients>, 15/12/2016
- [19] GUERFI ABABSA Souhila, « Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D », Université d'Evry Val d'Essonne, octobre 2008.
- [20] Mathias Ferraton. Réalisation d'un système d'acquisition 3D pour des objets transparents, Université de Bourgogne, 2010.
- [21] Système de Reconnaissance de Visage, université de Biskra, Septembre 2012.
- [22] Ayet SHAIK, « Reconnaissance d'objets 3D par points d'intérêt », l'École nationale supérieure des mines, Paris Mars 2013.
- [23] Volker Blanz et Thomas Vetter, Face Recognition Based on Fitting a 3D Morphable Model, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, volume: 25, Septembre 2003.
- [24] Analyse en Composantes Principales (ACP), <https://www.xlstat.com/fr/solutions/fonctionnalites/analyse-en-composantes-principales-acp>, 30/12/2018.
- [25] IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume : 14, Février 1992.
- [26] Lahoucine Ballihi, « Biométrie faciale 3D par apprentissage des caractéristiques géométriques : Application à la reconnaissance des visages et à la classification du genre », Université des Sciences et Technologie, Lille Mai 2012.
- [27] MESROUA Djamel et REBOUH Syphax, « Reconnaissance faciale par la télésurveillance », université Abderahmane Mira, Bejaia 2017.
- [28] OUAMANE Hanane, « Identification de reconnaissance faciale avec des expressions », Biskra Juin 2012.

[Bibliographie]

[29] HEDOUCHE Mourad, « RECONNAISSANCE DE VISAGE 3D BASÉE SUR LA LOCALISATION DU NEZ », Université Mohamed Khider, Biskra Juin 2013.

[30] Khefif Bouchra « Mise au point d'une application de reconnaissance faciale », Mémoire de fin d'études pour l'obtention du diplôme de Master en Informatique, Option : Réseaux et systèmes distribués (R.S.D.), Université Abou Bakr Belkaid – Tlemcen Faculté des Sciences, 28 novembre 2013.

[31] Pierre Guillot, « La méthode Viola et Jones et son implémentation dans OpenCV », Mars 2003.

[32] Jean-Philippe Vandeborre, « Analyse et reconnaissance de maillages 3D », Laboratoire d'informatique fondamentale, Lille Avril 2014.

[33] Hossen et al. « Face Detection by Using OpenCVs Viola-Jones Algorithm based on coding eyes », Iraqi Journal de Science, Vol. 58, Novembre 2017.

[34] Kristijan Ivancic « Traditional Face Detection With Python », 27 Février 2019.

[35] Abdelsamed Gorine et C.M Bentaouza « Reconnaissance faciale en utilisant le Deepface », Mémoire de Master en Informatique, Option : Ingénieries des Systèmes D'informations, Université Abdelhamid Ibn Badis, Mostaganem, Juin 2017.

[36] IEEE Int Conf Autom Face Gesture Recognit Workshops, « Dense 3D Face Alignment from 2D Videos in Real-Time », 2015 May.

[37] <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5082166-quest-ce-quun-reseau-de-neurones-convolutif-ou-cnn>, 16/04/2019.

[38] BEGGARI Samiha et KHAMRA Khaoula « Système de reconnaissance de visage par un réseau de neurone convolutionnel CNN », Mémoire de Master en Informatique, Option : Ingénieries des Systèmes D'informations, Université Kasdi Merbah, Ouargla, Mostaganem, Juillet 2017.

[39] <https://docs.python.org/fr/3/tutorial/index.html>, 23/05/2019.

[40] <https://tcuvelier.developpez.com/tutoriels/pyqt/bases/#LVI>, 26/05/2019.

[41] P. Lemaire¹, P. Szeptycki¹, M. Ardabilian¹, L. Chen, « Reconnaissance de visages en 3D orientée région », Ecole Centrale de Lyon, Département Math-Info, 2014 France.

[42] Mokri Mohammed Zakaria, « Classification des images avec les réseaux de neurones Convolutionnels », Université Abou Bakr Belkaid Tlemcen, Juillet 2017.

[Bibliographie]

[43] <http://www.statsoft.fr/concepts-statistiques/analyse-en-composantes-principales/analyse-en-composantes-principales.htm#classification>, 29/05/2019.

[44] <https://towardsdatascience.com/deep-learning-with-tensorflow-part-1-b19ce7803428>, 02/06/2019