

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Cheikh Larbi Tbessi –Tébessa-
Faculté des Sciences de l'Ingénieur
Département Informatique
Ecole Doctorale Informatique d'Est (Annaba)

N° D'ORDRE :
SERIE :



MEMOIRE DE MAGISTER
EN INFORMATIQUE
Option: Intelligence Artificielle (IA)

**Fusion des données pour un raisonneur par région
pour l'identification du visage humain**

Présenté par : Mr DJELLAB Issam

Dirigé par : Pr. Mohamed Tayeb LASKRI.

Soutenu le/2009

Devant le Jury composé de:

Président:	Dr. BOUDOUR Rachid	Maître de conférences	Université d'Annaba
Rapporteur:	Pr. Mohamed Tayeb LASKRI	Professeur	Université d'Annaba
Examineurs:	Dr KAZAR Okba	Maître de conférences	Université de Biskra
	Dr KIMOUR Tahar	Maître de conférences	Université d'Annaba

ملخص

إن التعرف الآلي على وجه الإنسان يبقى تحديا علميا دائم البحث، وهذا يرجع أساسا إلى طبيعة الوجه الإنساني الغني جدا بالتعبير الحسية و الانفعالية دائمة التغير و إلى طبيعة المحيط، التقدم في السن و وجود الشعر... ما يؤدي إلى جعل مهمة التعرف الآلي أكثر صعوبة. من هنا ظهرت عدة محاولات لحل هذه الإشكالية. لكن وبالرغم من التقنية العالية و التطور الملحوظ، لا توجد طريقة واحدة تلبى كل التغيرات التي قد تطرأ على الوجه. لذا، و محاولة لحصر أكبر عدد ممكن من هذه التغيرات، ظهرت بحوث جديدة ترمي إلى استعمال عدة مصنفات ثم دمج نتائجها بهدف الرفع من قدرة التعرف الآلي على الوجه.

العمل المقترح في هذه الورقة يدخل تحت لواء التعرف الآلي على وجه الإنسان، حيث يستند ترميز الخصائص إلى مبدأ لحساب المسافات، حيث انه و انطلاقا من صورة تحتوي على وجه نحاول أولا تحديد مجموعة من النقاط الأساسية التي تركز عليها صيغ حساب المسافات، هذه الأخيرة هي الخصائص التي تتيح لنا تحديد هوية الشخص، فباستخدام نظامنا نحدد هوية الفرد الموجود في الصورة أو نقول انه غير معروف، و لكي يتم التعرف على هوية الأشخاص استخدمنا طريقة تدعى أقرب ك نقطة مجاورة (KPPV) ، ولذلك يتم استخدام المسافة التي تحسب من خلال مجموع القيم المطلقة و التي تعرف أيضا باسم مسافة مينكوفسكي درجة 1 أو مسافة سيتي بلوك ، وهو ما يعادل مسافات أخرى (مثلا المسافة الإقليدية،...) ولكنها تتطلب حسابا أقل. إن فكرة دمج البيانات المقترحة تعتمد على انه يجب استخدام عدة طرق لحساب أقرب ك نقطة مجاورة وذلك لتحسين التعرف على هوية الأشخاص، وهذا يعني تطوير قائمة KPPV لكل طريقة (KPPV مع جميع الوجه، KPPV لكل منطقة في الوجه، KPPV للعيون والأنف، KPPV للعيون مع الفم، KPPV للأنف مع الفم). كل صور مجموعة التدريب تم تعريضها لمجموعة من العمليات (تعديل الحجم، النقر) بهدف استخراج مجموعة الخصائص التي تمثل كل وجه، بحيث أن خصائص الوجه هي عبارة عن المسافات بين النقاط الأساسية التي تم اختيارها يدويا.

الكلمات الدالة: التعرف على وجه الإنسان، أقرب ك نقطة مجاورة، التعرف على أشخاص.

Résumé

La reconnaissance automatique des visages humains demeure toujours un réel challenge dans le domaine de la recherche, ceci est dû essentiellement à la nature même de cette biométrie qu'est le visage très riche en informations sociales, et fortement dynamique à la fois à cause des variations de l'environnement, le vieillissement, la pilosité....chose qui rend la tâche de reconnaissance plus complexe; De ce fait, une multitude de méthodes ont été déployées pour résoudre ce problème, mais malgré la pléthore de techniques élaborées et les énormes progrès atteints, il n'existe pas une méthode qui soit à elle seule invariante à toutes les variations que peut subir ce stimulus. Pour essayer de cerner le maximum de ces changements, une nouvelle tendance a vu le jour visant à faire participer plusieurs classifieurs, et à combiner leurs raisonnements et leurs résultats en vue d'augmenter le taux de reconnaissance.

Le travail proposé dans ce mémoire s'inscrit dans le cadre de la reconnaissance des visages humains. Le codage des caractéristiques se base sur un principe d'extraction de distances. Autrement dit, ayant une image contenant un visage, d'abord nous cherchons à localiser des points d'intérêts sur lesquels reposent les formules de calcul des distances. Ces dernières représentent les caractéristiques permettant d'identifier une personne. Selon notre système, nous attribuons une identité à un individu ou nous le déclarons inconnu. La méthode de reconnaissance adoptée est celle des k plus proches voisins (KPPV). La distance utilisée est celle de la somme des valeurs absolues appelée aussi distance de Minkowski de degré 1 ou distance City-Block, elle est équivalente aux autres distances (ex. la distance euclidienne,...) mais elle nécessite moins de calculs (à ne pas confondre avec les distances AD, BC,...,SM), Avec l'approche de fusion des données proposée, plusieurs méthodes de calcul des k-ppv sont également mises au point pour optimiser la reconnaissance, cela veut dire une liste de k-ppv pour chaque méthode (k-ppv avec tout le visage, k-ppv pour chaque région du visage, k-ppv pour les yeux avec le nez, k-ppv pour les yeux avec la bouche, k-ppv pour le nez avec la bouche). Les images de la base d'apprentissage ont été sujet à une série d'opérations (normalisation, clique) afin d'extraire le vecteur des caractéristique représentant chaque visage.

Mots clés : reconnaissance des visages humains ; Identification de personnes ; Vecteur caractéristique ; Méthode des K-PPV.

Abstract

Automatic human face recognition remains a real challenge in the field of scientific research, because of nature of this biometric called human face, very rich of social information and highly dynamic, besides the variation environment, aging and hair... anything makes recognition more complex. For those reasons, multiple methods were deployed to resolve this problem. But in spite of the huge progress attained and, and the high technology used, there is no one method which could be told invariant to all the variations inflected by this stimulus. To surround all these variations, a next tendency was born, which aims to make participating many classifiers and combine their results with a view to enhance the recognition rate.

The work proposed in this dissertation constitutes a part of human faces recognition. The coding of features is based on a principle of distance extraction. In other words, having an image containing a face, first we try to locate points of interest underlying the formulas for calculating distances. The latter represents the characteristics that identify a person. In our system, we assign an identity to an individual or we declare him unknown. The recognition method adopted is that of k-nearest neighbors (KNN). The distance used is the sum of absolute values, also known as Minkowski distance of 1st degree or City-Block distance, it is equivalent to other distances (e.g. Euclidean distance...) but requires less computation (Not to be confused with the distance AD, BC... SM). However, along with the approach of data fusion proposed, there are several methods to calculate the k-NN developed to optimize the recognition i.e. a list of k-NN for each method (with k-NN entire face, k -NN for each facial region, k-NN for the eyes with the nose, k-NN for the eyes with the mouth, k-NN to the nose with the mouth). The images of the base of training were subjected to a series of operations (standardization, clicks) in order to extract the vector of the characteristic representing each face.

Key words: Human faces recognition; Identification of people; Characteristic vector; Method of the K-NN.

REMERCIEMENTS

Au terme de ce travail, je voudrai exprimer ma profonde gratitude envers Dieu le tout puissant qui, grâce à son aide, j'ai pu finir ce travail.

Je suis tout particulièrement reconnaissant à Monsieur M.T Laskri : Professeur à l'université de Badji Mokhtar, Annaba d'avoir accepté de diriger mon travail, pour la confiance qu'il m'a témoigné, pour ses précieux critiques, conseils. Toujours porteur de nouvelles idées, il a su donner à ce mémoire de thèse les bonnes orientations. Je voudrais également souligner son encouragement, son soutien constant et sa gentillesse incontestable, qu'il veuille trouver dans ce mémoire ma profonde gratitude et mon grand respect.

Mes remerciements les plus vifs sont adressés à Monsieur Boudour Rachid: Maître de Conférence à l'université de Badji Mokhtar, Annaba, qui a accepté de juger ce travail et d'en rédiger les rapports. Je remercie également Monsieur O. Kazar : Maître de conférences à l'université de Mohamed Khider, Biskra, ainsi que Monsieur M.T Kimour : Maître de Conférence à l'université de Badji Mokhtar, Annaba pour l'intérêt qu'ils portent à ce travail en acceptant d'en être les examinateurs.

Ma gratitude s'adresse aussi à tout les enseignants qui ont assuré ma formation du primaire jusqu'à L'université.

Mes remerciements les plus vifs s'adressent à Moussa Abdelhai, pour sa confiance, ses encouragements et sa gentillesse.

Enfin, je remercie toute ma famille, pour son soutien indéfectible, pour ses encouragements, et pour sa précieuse collaboration.

DEDICACE

Je dédie ce mémoire À mon père et à ma mère pour leur amour,
patience et leurs sacrifices,

À mon frère Belgacem,

À mes amies.

Liste des Figures

Figure 1. Schéma générale d'un processus de RDF.....	4
Figure.2 : Image d'un visage en noir et blanc (binaire).....	22
Figure. 3 : Image d'un visage en niveaux de gris.....	22
Figure. 4 : Modèle additive et soustractive des couleurs.....	23
Figure. 5 : Image couleur d'un visage.....	23
Figure.6 : Processus de traitement d'image.....	27
Figure 7 : Applications d'un filtre médian 3*3 sur une fenêtre 3*3.....	30
Figure. 8 : Applications d'un filtre minimum sur une fenêtre 3*3.....	30
Figure 9 : Marché des systèmes d'identification.....	38
Figure 10: Image d'iris	38
Figure 11: Empreinte digitale acquise par un capteur optique.....	39
Figure 12 : Deux types de minuties les plus utilisés dans la littérature.....	39
Figure 13: Reconnaissance de visage.....	40
Figure 14 : Schéma général du processus de Reconnaissance des visages.....	45
Figure 15 : exemple de la base (Faces94).....	68
Figure 16. Modèle géométrique du visage.....	69
Figure 17. Schéma bloc du système.....	71
Figure 18 : Distance globale.....	73
Figure 19 : Distance des yeux.....	73
Figure 20 : Distance des nez.....	74
Figure 21 : Distance des bouches.....	74
Figure 22 : Distance (yeux,nez).....	74
Figure 23 : Distance (yeux,bouche).....	74
Figure 24 : Distance (nez,bouche).....	75
Figure 25 page d'accueille de logiciel.....	76
Figure 26 : Ouvrir une image de la base.....	76
Figure 27 : Menu d'apprentissage.....	76
Figure 28 : Sélection des points caractéristiques pour l'apprentissage.....	76
Figure 29 : Menu de reconnaissance.....	77
Figure 30 : Sélection des points caractéristiques pour la reconnaissance.....	77
Figure 30 : Résultats obtenues en 10.1.....	78
Figure 31 : Résultats obtenues en 10.2.....	79

Table des Matières

Introduction Générale.....	01
----------------------------	----

Chapitre I : La Reconnaissance des Formes

1. introduction.....	03
2. Le schéma de la reconnaissance des formes	04
3. Les méthodes de la RdF	11
3.1. Méthodes statistiques.....	12
a. Décision bayésienne	12
b. Méthodes bayésiennes paramétriques.....	12
c. Méthodes bayésiennes non paramétriques.....	13
d. Méthodes paramétriques non bayésiennes.	13
e. Méthodes non paramétriques et non bayésiennes.....	13
• Les k-ppv (idée).....	13
ü Formulation de la méthode du plus proche voisin.....	14
ü Mise en place de la méthode.....	15
ü Avantages.....	16
ü Limites	16
f. Les réseaux connexionnistes.....	16
3.2. Méthodes structurelles.....	17
a. Comparaison dynamique de chaînes de symboles.....	18
b. Modèles de Markov.....	19
c. Modèle à états cachés.....	20

Chapitre II : L'Image sur Ordinateur

1. Introduction	21
2. L'image Numérique	21
2.1. L'image sur ordinateur.....	21
2.2. Les type d'images.....	22
2.2.1. Image noir et blanc (monochromes)	22
2.2.2. Images couleur	23
3. Représentation de l'image numérique.....	23
4. Caractéristiques d'une image numérique	24
4.1. Résolution d'image	24

4.2.	Image à niveaux de gris	24
4.3.	Image en couleurs	25
4.4.	Bruit	25
4.5.	Effets de bords	25
4.6.	Luminance	25
4.7.	Contraste.....	26
4.8.	Contours et textures	26
5.	Traitement et Analyse d'image.....	26
5.1.	Acquisition.....	27
5.2.	Filtrage.....	28
5.3.	Segmentation d'Image.....	31
5.3.1.	Segmentation par extraction de régions.. ..	31
5.3.2	Segmentation par extraction de contours :	31
5.4.	Interprétation	33
6.	Domaine d'application	33

Chapitre III : La biométrie et La reconnaissance des visages humains

1.	Définition	35
2.	Système biométrique.....	35
3.	Erreurs du système biométrique	36
4.	Applications de la biométrie	36
5.	Techniques biométriques	37
6.	L'Industrie de la Biométrie	37
7.	quelques Techniques biométriques.....	38
I.	La reconnaissance d'Iris	38
II.	Empreintes digitales	39
III.	Signature.....	40
IV.	La reconnaissance de Visage.....	40
IV.1.	Caractéristiques Visage	41
IV.1.	Définition et Particularités du visage :	41
IV.2	Pourquoi le visage ?	42
IV .3.	D'où vient la difficulté ?	43
IV.4.	Architecture Générale d'un Système de Reconnaissance de Visages.....	45

IV.5.	Détection des visages	48
IV.6.	Extraction des Paramètres du Visage.....	49
IV.6.1	L'intégrale de projection verticale et horizontale.....	50
IV.6.2	La méthode des moments invariants	50
IV.6.3	La transformée de symétrie généralisée.....	51
IV.7.	Les Méthodes de Reconnaissance des Visages.....	53
IV.7.1	Principe.....	53
IV.7.2.	Les Méthodes Existantes.....	53
IV.7.2.1.	Les Méthodes Géométriques.....	54
IV.7.2.2.	Les Méthodes Globales.....	56
IV.7.2.2.1	Les Visages Propres (ACP).....	56
IV.7.2.2.2	Les Réseaux de Neurones.....	57
IV.7.2.2.3	Les Templates Matching(Appariement de modèles).....	59
IV.7.2.2.4	Les Machines à Vecteurs de Support SVM.....	59
IV.7.2.2.5	Les Modèles de Markov Cachés HMM.....	61
IV.7.2.2.6	Les Systèmes Multi-classifieurs.....	62
IV.7.3	Avantages et Limites des méthodes de reconnaissance des visages.....	63
IV.8.	principaux domaines d'utilisations.....	63
9.	Avantage de la biométrie.....	64
10.	Conclusion.....	64

Chapitre IV :

1.	Introduction.....	65
2.	Motivation.....	66
2.1.	Pourquoi la reconnaissance des visages.....	66
2.2.	Pourquoi les k plus proches voisins.....	66
3.	présentation de Notre Système.....	67
4.	La base des images utilisée	67
5.	Système de reconnaissance de visages.....	68
6.	Architecture de Système	70
7.	Un système d'identification.....	72
8.	Méthode de reconnaissance.....	72
8.1.	Détermination des k-ppv d'après tout le visage.....	73
8.2.	Détermination des k-ppv pour chaque région du visage.....	73

8.2.1. Détermination des k-ppv pour les yeux.....	73
8.2.2. Détermination des k-ppv pour le nez.....	73
8.2.3. Détermination des k-ppv pour la bouche.....	74
8.3. Détermination des k-ppv pour les yeux avec le nez.....	74
8.4. Détermination des k-ppv pour les yeux avec la bouche.....	74
8.5. Détermination des k-ppv pour le nez avec la bouche.....	75
9. Etude comparative.....	75
10. Présentation de l'application	76
11. Résultats et discussions.....	80
Conclusion générale.....	81

Introduction Générale

La vision joue un rôle primordial dans notre vie, en effet c'est la meilleure façon de comprendre le monde extérieur. L'homme est considéré comme un expert dans la vision, il peut facilement voir et reconnaître les objets et les personnes qu'il voit et avec une très grande précision. Cette capacité à permet à plusieurs chercheurs de s'en inspirer pour réaliser des systèmes et créer des machines capables de " voir ", d'où vient la naissance d'un nouveau domaine de l'informatique qui est la vision par ordinateur.

La vision par ordinateur est devenue depuis plusieurs années un domaine très actif, ces recherches se basent sur la simulation du processus de la vision humaine pour la création des systèmes de vision automatiques. Elle regroupe plusieurs sous domaines parmi lesquels celui de biométrie, dont fait partie la reconnaissance des visages.

Un système automatique de reconnaissance de visages est un système biométrique utilisant le visage à des fins d'identification et/ou de vérification de personnes à partir de leurs images de visages fixes ou de séquences vidéo en comparant les caractéristiques de cet individu avec celles stockées dans une base de référence. Comparé à d'autres systèmes biométriques, il a l'avantage, d'être non intrusif, effectuant l'identification d'une façon naturelle et acceptable par les utilisateurs.

Une des techniques propose Le visage numérisé comme stratégie de reconnaissance. Identifier un individu par son visage, nécessite le passage par plusieurs étapes: prise de vue, extraction de son visage, codage de ses caractéristiques et identification par comparaison avec le contenu d'une base contenant plusieurs centaines d'images du visage humain.

Comme tout système de reconnaissance de formes (RDF), un système de reconnaissance de visages (RDV) comporte une étape de classification pour laquelle plusieurs classificateurs ont été adoptés. Les plus utilisés, dès les premières approches et jusqu'aux plus récentes, sont ceux basés sur la règle du plus proche voisin avec la distance euclidienne ; à cause de sa simplicité. Cependant, nous devons noter l'influence de la distance choisie sur la performance du système. Les réseaux de neurones ont aussi attiré les chercheurs motivés par leur pouvoir de généralisation et leur auto adaptabilité. Ils ont été appliqués à des fins de détection de visages, de reconnaissance d'expression faciale, de détection, localisation et reconnaissance de visages et vérification de visages. Historiquement, l'intérêt à la reconnaissance de visages remonte à 1888 avec les travaux

de Galton à des fins d'identification de personnes à partir de leur vues de profil. Les recherches concernant la reconnaissance automatique de visages n'ont vu le jour qu'en 1966 avec le premier système semi automatique de Bledsoe. Le premier système entièrement automatique, proposé par T.Kanade date de 1973... En 1991, la publication de l'article "Eigenfaces for Recognition" fut un tournant dans le domaine de la reconnaissance de visages ; depuis les articles et les publications nécessitent d'apparaître en s'inspirant des eigenfaces.

Le travail proposé dans ce mémoire entre dans le cadre de la reconnaissance de visages humain. Le codage des caractéristiques se base sur un principe d'extraction de distances. C'est à dire, partant d'une image qui contient un visage on cherche tout d'abord à localiser des points d'intérêts sur les quels reposent les formules de calcul des distances. Ces derniers représentent les caractéristiques qui nous permettront d'identifier une personne. Avec notre système on attribue une identité à un individu ou on le déclare inconnu. La méthode de reconnaissance adoptée est celle des k plus proches voisins (KPPV). La distance utilisée est celle de la somme des valeurs absolues appelée aussi distance de Minkowski de degré 1 ou distance City-Block, elle est équivalente, aux autres distances (e.g. la distance euclidienne,...) mais elle nécessite moins de calculs. Avec l'approche de fusion des données proposée, plusieurs méthodes de calcul des k-ppv sont mises au point pour optimiser la reconnaissance, c'est-à-dire une liste de k-ppv pour chaque méthode(k-ppv avec tout le visage, k-ppv pour chaque région du visage, k-ppv pour les yeux avec le nez, k-ppv pour les yeux avec la bouche, k-ppv pour le nez avec la bouche).

Plan du mémoire :

Cette thèse est organisée comme suit :

- Ø Le chapitre I survole le concept de ***Reconnaissance des formes***.
- Ø Le chapitre II ***L'Image sur ordinateur***.
- Ø Le chapitre III présente la notion de ***la biométrie et reconnaissance des visages humains***.
- Ø Le chapitre IV fournit ***La conception et la réalisation de l'architecture proposée***.
- Ø Une ***conclusion générale*** marque le point final de ce travail.

Chapitre I

Reconnaissance des formes

1. Introduction

L'idée de construire des machines capables de simuler des êtres humains afin de les aider dans certaines tâches, voire de les remplacer, était antérieure aux ordinateurs. Leur apparition a permis d'étendre le spectre des tâches à simuler en ajoutant celles dont l'exécution relève de facultés mentales comme la perception et le raisonnement. [64]

Le problème que cherche à résoudre la reconnaissance des formes est d'associer une étiquette à une donnée qui peut se présenter sous forme d'une image ou d'un signal. Des données différentes peuvent recevoir la même étiquette, ces données sont les réalisations ou les exemplaires de la classe identifiée par l'étiquette. Par exemple, le son /a/ prononcé par différents locuteurs conduit à des signaux différents mais ces différences ne sont pas significatives du point de vue de l'identification du son, ces signaux sont des réalisations de la classe /a/. De même, l'écriture manuscrite du caractère A varie d'un scripteur à l'autre mais le lecteur identifiera le caractère A pour chacune de ces réalisations.

Des méthodes générales ont été développées en reconnaissance des formes pour extraire automatiquement des informations des données sensibles afin de caractériser les classes de formes (apprentissage) et d'assigner automatiquement des données à ces classes (reconnaissance). La mise en œuvre de ces méthodes générales pour des problèmes particuliers amène à introduire la notion de processus de reconnaissance qui pose la question de l'intégration des méthodes de la reconnaissance de formes dans un système qui a pour but de reconnaître des formes.

Parallèlement aux travaux sur les méthodes de reconnaissance, se développaient le traitement d'image, la vision par ordinateur, et le traitement de la parole. Ces domaines ont focalisé le problème de la reconnaissance sur des données spécifiques, mais par ailleurs ils ont permis de situer la reconnaissance dans un processus plus vaste d'interprétation d'image ou de compréhension de la parole impliquant des niveaux de perception et des connaissances propres au domaine.

2. Le schéma de la reconnaissance des formes

On s'appuie sur le schéma classique d'un processus de reconnaissance de formes pour décrire les principaux traitements à effectuer et leurs objectifs.

Buts des étapes du schéma :

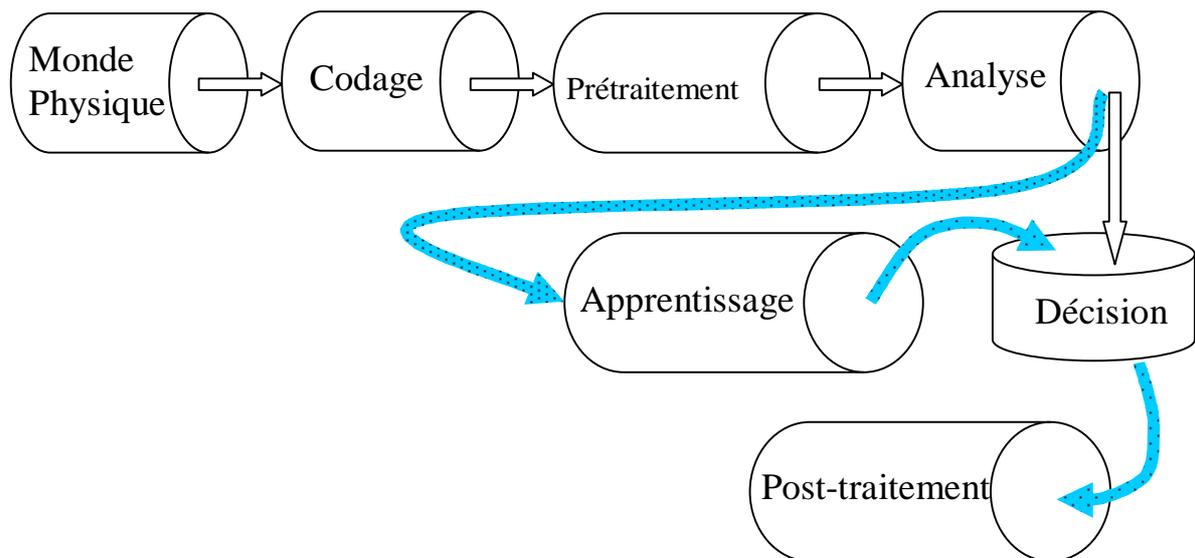


Figure 1. Schéma générale d'un processus de RDF

§ Numérisation (codage) : obtenir une représentation des données à traiter qui soit manipulable en machine.

§ Prétraitement : élimination des bruits, normalisation, re-échantillonnage, amélioration des contrastes, etc.

§ Calcul des représentations (Analyse) : obtenir une représentation des données compatible avec les outils d'apprentissage et de décision utilisés.

§ Apprentissage : à partir d'un ensemble d'exemplaires, construire une représentation des classes.

§ Décisions : assigner une forme inconnue à une classe.

§ Post-traitement : valider les décisions de l'analyse sur la base de connaissances (du domaine).

Dans la pratique, un système de reconnaissance des formes s'éloigne souvent de ce schéma. Des traitements en amont sont souvent nécessaires pour isoler la forme à reconnaître de son contexte, ce qui en soit est un problème de reconnaissance (segmentation forme/fond, délimitation d'une forme dans un ensemble). Des traitements ultérieurs sont aussi utiles pour valider les décisions et éventuellement les remettre en cause.

Le caractère séquentiel du processus de reconnaissance, tel qu'il est présenté dans ce schéma, n'est pas toujours la meilleure option. Ainsi, une erreur dans la segmentation de la forme à reconnaître augmente forcément le risque d'une mauvaise reconnaissance. Il est possible d'introduire une boucle dans le processus, remettant en cause la segmentation après analyse des résultats de la reconnaissance. Il est aussi possible de fusionner segmentation et reconnaissance. C'est le cas en écriture où la segmentation en lettres d'un mot cursif peut être couplée à la reconnaissance des suites de lettres formant un mot lexicalement valide (même principe pour la segmentation d'un signal de parole en phonèmes). Ce schéma illustre en fait l'immersion du noyau dur de la reconnaissance de formes (les méthodes d'apprentissage et d'assignation d'une donnée à une classe) dans un système qui, partant d'un signal brut, va permettre de résoudre un problème de reconnaissance.

- **Décisions** [64]

Le problème de l'assignation d'une donnée à une classe peut se poser de différentes manières. On distingue le cas où des données étiquetées par un **maître** sont disponibles du cas où aucun moyen pour connaître les classes d'appartenance des données n'existe. Dans ce dernier cas, le traitement automatique doit faire apparaître des groupements de données. Il faut spécifier les représentations et une mesure de distance dans l'espace de représentation qui soient les mieux appropriées pour mettre ces groupements en évidence. La validation de ces groupements est un problème en soit. Elle peut être faite à posteriori, à l'issu d'une étude de confirmation menée sur le domaine d'application et guidée par les résultats de la classification de données.

Dans le cas de la simulation d'une reconnaissance humaine, le maître existe et la décision peut prendre deux formes : la **discrimination** ou la **caractérisation**. Pour la discrimination, il s'agit d'assigner une forme inconnue à une des classes possibles (exemples de la reconnaissance des caractères, des phonèmes). Dans le cas de la caractérisation, il s'agit de décider si une forme possède les caractéristiques d'une classe ou non (exemple : décider si un électrocardiogramme appartient à la classe "normal", rechercher un mot clé dans une séquence parlé).

L'apprentissage des représentations des classes ne se fait pas de la même manière dans les deux cas. Pour la discrimination, on recherche ce qui fait la différence entre les classes, les caractères discriminants. On utilise les exemplaires des classes j comme contre-exemples de la classe i , ce qui permet de mieux cerner ce qui fait la différence entre les classes. Dans le cas de la caractérisation, les contre-exemples de la classe unique à caractériser correspondent à tout ce qui n'est pas un exemplaire de cette classe mais la variété des contre-exemples est telle qu'il est difficile d'en extraire une information pertinente pour préciser la classe à caractériser. Par exemple, l'ensemble des mots de la langue parlé sont les contre-exemples de la classe formée par les différentes productions d'un mot cible. Caractériser ce mot cible, plus exactement sa classe, consiste à rechercher les invariants, c'est-à-dire ce qui reste stable quand on considère un ensemble d'exemplaires de la classe.

La validation de ces méthodes se fait par comparaison des résultats de la reconnaissance automatique aux étiquettes données par le maître. On en tire donc les taux de **reconnaissance** et les taux d'**erreur**. On peut aussi avoir des taux de **rejet** qui correspondent à la décision de ne pas classer la forme.

Le système en évaluant un critère de décision peut assigner une forme à une classe mais il peut aussi déterminer avec quelle confiance il effectue cette décision. Si le critère de décision prend des valeurs très proches pour plusieurs classes, la confiance dans la décision est faible.

Le rejet est très important dans les applications, il va être modulé en fonction d'impératifs de risque et de cout. Par exemple, il est très couteux de faire des erreurs sur la lecture des adresses postales et encore plus sur les montants des chèques, on préférera donc augmenter les taux de rejet quitte à baisser les taux de reconnaissance. Par contre, la surveillance médicale nécessite de fort taux de reconnaissance et surtout de ne pas prendre le risque de rater un événement critique en confondant des données anormales avec des données normales. La

encore, il faudra ajuster les seuils de décisions en fonction du risque et préférer des alarmes intempestives à des erreurs de non détection de cas critiques.

La décision finale est en général le résultat de plusieurs **décisions intermédiaires** qui peuvent être organisées hiérarchiquement. On distingue une hiérarchie d'**affinement** où l'identité de la forme à reconnaître sera précisée à mesure que le processus se déroule. Une première étape de segmentation permet de décider qu'un sous-ensemble des données enregistrées constitue une des formes du domaine, les étapes suivantes consistent à mesurer et évaluer des caractéristiques de cette forme pour parcourir une hiérarchie de classes du sommet aux feuilles qui correspondent aux décisions finales. Par exemple, pour la reconnaissance de figures géométriques : la classe "figure fermée" se divise en "polygone" et "conique" qui se divise en "cercle" et "ellipse". On peut aussi avoir une hiérarchie de décisions définie par la taille du **contexte** qui augmente à mesure que le processus se déroule. Par exemple pour la reconnaissance de mots écrits : reconnaissance des composantes primitives du tracé, reconnaissance des caractères, reconnaissance des mots.

- **Prétraitement** [64]

Les **données brutes** issues des capteurs sont les représentations initiales des données à partir desquelles des traitements permettent de construire celles qui seront utilisées pour la reconnaissance. Les données brutes sont bruitées, elles contiennent des informations parasites, et elles n'explicitent pas les informations utiles pour la reconnaissance. Par exemple, l'information sur la forme des objets à reconnaître est contenue dans le tableau de pixels résultant de la numérisation d'une image mais des traitements seront nécessaires pour, par exemple, séparer la forme du fond et suivre son contour.

Pour les prétraitements, le concepteur s'aide des connaissances qu'il possède sur les capteurs, les types données, le problème posé et les méthodes d'apprentissage et de reconnaissance qu'il utilisera. Les prétraitements sont utiles pour éliminer des bruits qui peuvent être dus au capteur ou à des interférences avec d'autres sources de signaux (la parole en milieu sonore, l'encre du verso qui traverse le papier et dont la trace est visible sur la feuille du manuscrit, les fonds imagés des chèques, etc.).

Le traitement du signal et de l'image sont les principales sources pour les méthodes de prétraitement : filtrage (des bruits hautes fréquences, par exemple), amélioration des contrastes,

extraction de contours ou de squelettes, modélisation du signal temporel, extraction des primitives, etc.

Les **prétraitements** et l'**apprentissage** ont des buts communs qui consistent à éliminer les informations parasites et à conserver les informations pertinentes pour la reconnaissance. La limite entre pré-traitement et apprentissage relève des possibilités des méthodes d'apprentissage utilisées.

- **Représentation**

Le système de reconnaissance de formes doit posséder des représentation des formes à classer et des classes. [65]

Les types de représentation sont déterminés par les méthodes utilisées. Les contenus des représentations sont déterminées par les buts de la reconnaissance.

Certains types de représentation sont mieux appropriés à exprimer certains contenus. Par exemple, une variabilité intra-classe qui se présente comme une répartition aléatoire des caractéristiques de la forme autour de valeurs fortement probables, sera bien représentée par des modèles statistiques. Par contre, les structures (les relations entre composantes d'une forme) peuvent plus facilement s'exprimer par des représentations structurelles.

On appelle **caractéristique**(ou descripteur) une information qui peut être mesurée sur la donnée à reconnaître. Par exemple : l'amplitude moyenne d'un signal sur une fenêtre temporelle, l'énergie dans une bande de fréquence, le rapport hauteur sur largeur d'un caractère manuscrit, le niveau de gris moyen d'une zone d'image etc.

On appelle **primitive** une composante élémentaire d'une forme, les primitives ne sont pas décomposables. Par exemple : un segment de droite, une boucle ...

- **Représentation d'un exemplaire :**

- Un vecteur de dimensions n . Chaque dimension correspond à la mesure d'une caractéristique sur l'exemplaire. Chaque exemplaire est alors un point de R^n .
- Une organisation de primitives sous forme de séquences, de graphes ou de règles de composition.

-
-
- Un ensemble de propriétés, une description logique.
 - Des fonctions, équations paramétriques des contours etc.
 - **Représentation des classes**

Dans le cas où l'on dispose d'un ensemble d'exemplaires étiquetés, on distingue les démarches qui pratiquent un apprentissage supervisé pour induire la représentation des classes de celles qui ne pratiquent pas d'apprentissage. Dans tous les cas, on recherche des représentations des classes qui permettront d'obtenir les meilleures performances en reconnaissance.

L'**apprentissage** a pour but de **généraliser** l'information portée par les exemplaires de l'ensemble d'apprentissage pour faire émerger les représentations des classes. Ces représentations doivent permettre de s'affranchir de la variabilité des exemplaires d'une classe lors de la prise de décision. La validation de l'apprentissage est faite par les taux de reconnaissance obtenus sur un ensemble test constitué d'exemplaires différents de ceux de l'ensemble d'apprentissage. Pour les démarches qui excluent cet apprentissage, le critère de décision (une mesure de ressemblance) devra supporter seul la variabilité intra-classe.

Le type de représentation choisi pour les exemplaires (un vecteur de R_n , une composition de symboles ...) va conditionner le type de représentation des classes (un sous-espace de R_n , une grammaire ...).

- **Représentation d'une classe :**

En extension :

- La représentation de un ou plusieurs exemplaires qui seront considérés comme des prototypes de la classe.
- La représentation d'un ensemble d'exemplaires de la classe qui couvre l'espace des réalisations possibles pour cette classe. [\[65\]](#)

En intension :

- Un sous espace de R_n .
- Un langage, un graphe.
- Une description sous forme de composition de propriétés.

Il est aussi possible de représenter la classes par un ou plusieurs prototypes qui ne correspondent pas à des exemplaires réellement observés mais sont construits à partir d'exemplaires de la classe, ils les "résumant".

- **Ressemblance**

Ce qui fait problème pour reconnaître des formes comme étant des exemplaires d'une classe est d'une part la **variabilité intra-classe** et d'autre part la **séparabilité** des classes. Les exemplaires d'une classe ne sont pas identiques, ils se ressemblent, alors que les exemplaires appartenant à des classes différentes ne se ressemblent pas. Le problème sera donc de formaliser cette ressemblance.

On peut supposer que la ressemblance des exemplaires d'une classe est due à la présence d'invariants caractéristiques de la classe. Par exemple, les rectangles ont toujours quatre cotés qui forment des angles droits. En fait, les réalisations à main levée de rectangles s'éloignent fortement de cette définition. Ce qui conditionne la décision est le fait que la figure ressemble d'avantage à un rectangle qu'à une autre figure géométrique, elle en est plus proche au sens d'une métrique à définir.

La répartition des exemplaires dans l'espace de représentation fait rarement apparaître une parfaite séparation entre les différentes classes, ce qui provoque une ambiguïté dans les décisions. On peut associer une valeur à la décision pour indiquer avec quelle confiance elle a été prise.

Dans le cas où les classes sont représentées en extension, la ressemblance peut être mesurée :

- par la distance d'un exemplaire inconnue à un prototype, c'est l'appariement de formes (Pattern Matching). Méthodes : Programmation dynamique pour la comparaison de chaînes, de graphes, masques élastiques ...
- par la distance d'un exemplaire inconnu à des exemplaires dont on connaît la classe d'appartenance Méthode des k-plus-proches-voisins.

Dans le cas où les classes sont représentées par des sous espaces de R_n , des langages, des descriptions logiques, la ressemblance est donnée par la position de l'exemplaire inconnu par rapport aux espaces définis par les représentations des classes. Méthodes : théories statistiques

de la décision pour un espace R_n , analyse syntaxique pour des classes représentées par des langages ...

3. Les méthodes de la RdF [63]

Les méthodes de la RdF sont souvent regroupées en grandes classes identifiées par : *statistique, syntaxique, structurelle, hybride (une combinaison des autres)*.

A ces classes correspondent différentes manières de représenter les exemplaires et les classes et différentes méthodes pour l'apprentissage et la reconnaissance. Mais elles correspondent aussi à différentes façons d'aborder le problème de la reconnaissance de formes. On s'intéresse ici à la simulation d'un être humain dans une tâche de reconnaissance, ce qui ne couvre pas l'ensemble des problèmes traités par la RdF.

La première façon d'aborder la simulation d'une reconnaissance humaine consiste en un traitement entièrement automatisé où le "maître" à simuler étiquette les réalisations qui seront utilisées pour l'apprentissage et pour la validation de la reconnaissance. Les méthodes statistiques et syntaxiques répondent à cette démarche. La difficulté à apprendre automatiquement les descriptions qui caractérisent les classes a conduit à demander à l'expert (le maître devient l'expert) d'explicitier ces descriptions, d'aider à construire la représentation des classes et les critères de décision. C'est là que les méthodes structurelles interviennent. En fait la reconnaissance ne se ramène pas à une analyse de la forme, cette analyse est souvent un élément dans un processus de décision qui implique d'autres informations ou d'autres connaissances (l'âge de la personne sur laquelle un électrocardiogramme a été enregistré, l'heure à laquelle la photo aérienne a été prise etc.). Les systèmes à base de connaissances tentent de reproduire les conditions dans lesquelles l'expert prend sa décision. Il doit donc pouvoir raisonner à partir de ce qu'il voit (perception) et de ce qu'il sait (connaissance). Chaque nouvelle démarche dans l'évolution de la RdF ne signifie pas un remplacement des démarches et méthodes passées mais un enrichissement de l'ensemble des méthodes dans lequel il faudra trouver celles qui sont les mieux adaptées au problème traité.

3.1. Méthodes statistiques

On considère ici la situation où il existe un ensemble de réalisations étiquetées permettant un **apprentissage supervisé** (par opposition aux méthodes qui partent d'une connaissance nulle sur les étiquettes des réalisations et cherchent à les grouper sur des critères de ressemblance).

Les exemplaires des classes correspondent aux observations d'une variable aléatoire X . Chaque réalisation x est représentée par un **vecteur de \mathbf{R}^n** . Chaque composante du vecteur correspond à un descripteur. [63]

L'objectif est toujours d'assigner une réalisation inconnue à sa classe d'appartenance en **minimisant l'erreur de décision**. Ce problème peut être résolu de nombreuses manières, le choix de la méthode dépend en partie des connaissances a priori que l'on a sur les distributions de probabilités des exemplaires des classes.

Le cas Bayésien correspond à des critères de décision construits à partir des lois de probabilité. Le cas paramétrique correspond à une connaissance a priori sur les lois de probabilité dont il faut estimer les paramètres. On arrive ainsi à une classification des méthodes

a. Décision bayésienne

La décision qui minimise la probabilité d'erreur globale est celle qui associe à chaque point x de \mathbf{R}^n la classe dont la densité de probabilité en x est la plus forte.

Les classes : $\{w_i\}, i = 1 \dots s$

Règle de Bayes : $P(w_i | x) = \frac{P(x|w_i)P(w_i)}{P(x)}$ avec $p(x) = \sum_{i=1}^s p(x|w_i)P(w_i)$

Problèmes : estimer $P(w_i)$ et $p(x|w_i)$

b. Méthodes bayésiennes paramétriques :

Estimer les paramètres des lois de probabilité connues a priori (lois normales en général).
Estimateurs de la moyenne et de la matrice de covariance :

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{et} \quad \hat{\Sigma} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \hat{\mu})(x_i - \hat{\mu})^t$$

Pour des lois normales, la règle de Bayes se ramène à maximiser :

$$\frac{P(w_i)}{(2\pi)^{1/2}(\text{Det}\Sigma_i)^{1/2}} e^{-\frac{1}{2}(x-\hat{\mu}_i)^t \Sigma_i^{-1} (x-\hat{\mu}_i)}$$

soit :

$$-(x - \mu_i)^t \Sigma_i^{-1} - [\text{Log}(\text{Det}(\Sigma_i)) - 2\text{Log}P(w_i)]$$

Le premier terme est la distance de Mahalanobis, le terme entre crochet ne dépend pas de la réalisation x . Si les classes sont équiprobables, le terme contenant la probabilité des classes s'annule. Si les classes ont des covariances égales, le terme contenant la matrice de covariance s'annule.

c. Méthodes bayésiennes non paramétriques :

Pas de lois de probabilité connue a priori. L'espace est discrétisé et des estimateurs de loi (fenêtres de Parzen) sont utilisés.

Autres méthodes

d. Méthodes paramétriques non bayésiennes.

Construction des surfaces séparatrices des classes sans estimer les probabilités. La surface séparatrice a une forme connue, il faut estimer les paramètres qui optimisent la décision.

Catalogue :

- Algorithme du perceptron
- Algorithme de Ho Kashyap (gradient)
- Méthode de Fisher

e. Méthodes non paramétriques et non bayésiennes.

Recherche d'une partition de l'espace de représentation sans rechercher les lois de probabilités ni les surfaces de séparation. La méthode des k-plus-proches-voisins en est un exemple.

Les k-ppv (idée) : Dans l'espace de représentation R_n , les k-plus-proches-voisins de la réalisation inconnue définissent le sous-espace de décision. La décision se fait par un vote, la réalisation inconnue hérite de l'étiquette de la classe la plus représentée dans ce sous-espace.

La méthode du plus proche voisin est la méthode locale non paramétrique la plus simple. On suppose seulement que le voisinage qui influe sur la décision de classification du point inconnu se limite à une boule contenant seulement le plus proche voisin. La densité de probabilité de toutes les classes est nulle, à l'exception de celle affectée au plus proche voisin qui est naturellement choisie pour la décision. Cette méthode en apparence très simple est robuste. Son erreur est majorée par 2 fois l'erreur minimale obtenue par la classification bayésienne, dans les conditions asymptotiques. Elle permet de reconnaître les formes lorsque la distribution des classes n'est pas convexe. L'algorithme de base est facile à écrire et à programmer. En contrepartie, elle nécessite un grand volume de données d'apprentissage, ce qui veut dire un grand nombre de points d'apprentissage à stocker et à examiner, autrement dit des ressources mémoire élevées et un temps d'exécution important dans une implémentation brute de l'algorithme.

ü Formulation de la méthode du plus proche voisin

La mise en œuvre mathématique de la méthode des K plus proches voisins peut être formulé comme suit : [66]

Algorithme de classification par k-PPV

Paramètre : le nombre k de voisins

Contexte : un échantillon de L objets classés en $C = c_1, c_2, \dots, c_n$ classes

- 1: **pour** chaque objet O **faire**
- 2: transformer (exprimer) l'objet O en vecteur $t = (x_1, x_2, \dots, x_m)$
- 3: déterminer les k plus proches objets de l'objet O selon une métrique de distance
- 4: combiner les classes de ces k exemples en une classe c
- 5: **fin pour**

Sortie : l'Objet O associé à la classe c.

Les choix de la distance et du paramètre k sont primordiaux pour le bon fonctionnement de cette méthode. Les différents choix possibles pour la *distance* et pour le *mode de sélection* de la classe du cas présenté.

▼ Définition de la distance

Une distance est une application de $E \times E$ dans R^+ telle que les propriétés suivantes soient vérifiées [67] : pour tous les points a, b, c de l'espace.

$$d(a, b) \geq 0$$

$$d(a, b) = 0, a = b$$

$$d(a, b) = d(b, a)$$

$$d(a, b) \leq d(a, c) + d(c, b)$$

La distance entre un objet et ses voisins se fait via une métrique de distance. Cette métrique peut être la métrique de Minkowski de degré (n)

$$d_n(a, b) = \sqrt[n]{\left(\sum_i |a_i - b_i|^n\right)}$$

Selon la valeur de n, on retrouve plusieurs distances connues :

1. Si $n = 1$ cette distance est la distance de Manhattan définie par

$$d_m(a, b) = \{|a_1 - b_1| + |a_2 - b_2| + \dots + |a_n - b_n|\}$$

2. Si $n = 2$ c'est la distance euclidienne définie par

$$d_e(a, b) = \sqrt{\left(\sum_i (a_i - b_i)^2\right)}$$

3. Si $n = \infty$ c'est la distance de Chebyshev définie par

$$d_c(a, b) = \max\{|a_1 - b_1|, |a_2 - b_2|, \dots, |a_n - b_n|\}$$

ü Mise en place de la méthode [06]

La méthode ne nécessite pas de phase d'apprentissage. Le modèle est constitué des trois éléments :

- 1) l'échantillon d'apprentissage.
- 2) la distance
- 3) la méthode de combinaison des voisins.

L'efficacité de la méthode dépend de ces trois éléments. Il faut choisir l'échantillon, c'est-à-dire les attributs pertinents pour la tâche de classification considérée et l'ensemble des enregistrements. Il faut veiller à disposer d'un nombre assez grand d'enregistrements par

rapport au nombre d'attributs et à ce que chacune des classes soit bien représentée dans l'échantillon choisi [09].

Il est possible d'optimiser la distance en faisant varier les paramètres et en estimant l'erreur en généralisation pour chacun des choix. L'estimation de l'erreur en généralisation se fait classiquement, soit avec un ensemble test, soit en validation croisée.

Le choix du nombre k de voisins peut, lui aussi, être déterminé par utilisation d'un ensemble test ou par validation croisée. Une heuristique fréquemment utilisée est de prendre k égal au nombre d'attributs plus 1. La méthode de vote ou d'estimation peut aussi être choisie par test ou validation croisée.

ü **Avantages :**

- Ø Simplicité de la mise en œuvre.
- Ø La méthode ne nécessite pas de phase d'apprentissage; c'est l'échantillon d'apprentissage, associé à une fonction de distance et à une fonction de choix de la classe en fonction des classes des voisins les plus proches, qui constitue le modèle [06]
- Ø Il n'y a pas d'étape d'apprentissage plutôt la construction d'un modèle.

ü **Limites :** [68]

- Ø Les algorithmes de type plus proche voisin sont longs en phase de généralisation, puisqu'ils sauvegardent tous les exemples de l'échantillon d'apprentissage ;
- Ø Ils sont sensibles au bruit sur les variables prédictives ;
- Ø Ils sont sensibles au choix de la fonction de similarité de l'algorithme.

f. Les réseaux connexionnistes

Les méthodes connexionnistes constituent un domaine de recherche à part entière mais elles se sont imposées en reconnaissance de formes. Elles y trouvent matière à développer et à valider de multiples variantes de réseaux. Elles permettent d'étendre les techniques de séparation linéaire (le cas du perceptron) à des séparations non-linéaires.

L'idée des réseaux connexionnistes a été inspirée par la structure neurophysiologique du cerveau. Un neurone formel est une unité d'un système organisé en réseau par des connexions entre unités.

Quelques notions :

Un neurone formel (ou cellule) va recevoir des signaux provenant d'autres cellules du réseau et réagir en produisant une sortie qui sera reçue par des cellules du réseau.

L'état d'activation d'une cellule j est une somme pondérée des signaux d'entrée émis par un ensemble de cellule i . Les poids w_{ij} caractérisent l'influence d'un signal émis par une cellule i sur la cellule j . La sortie y_j produite par j est une fonction de son état d'activation, soit :

$$y_j = f \left(\sum_{i=1}^n W_{ij} X_i \right)$$

C'est par l'intermédiaire de la fonction f (le plus souvent une sigmoïde) qu'est introduite la notion de non-linéarité.

Dans les réseaux à couches, les cellules sont de trois types :

- les cellules d'entrée qui sont associés aux données, il y a analogie entre les composantes un vecteur de représentation et les cellules de la couche d'entrée,
- les cellules de sortie, chacune d'elle étant la représentante d'une classe de formes,
- les cellules cachées qui sont entre la couche d'entrée et de sortie.

L'apprentissage à partir d'exemples consiste à faire évoluer les poids w_{ij} de manière à minimiser l'erreur entre les sorties désirées et les sorties observées. C'est un apprentissage incrémental basé sur la technique du gradient.

3.2. Méthodes structurelles

On peut dire, pour dégager l'essentiel de la différence entre méthodes statistiques et méthodes structurelles, que d'une part une donnée est représentée par des vecteurs de caractéristiques et que d'autre part elle est représentée par une description sous forme de composantes et de relations entre ces composantes.

La notion de structure, quoique sujette à de nombreuses définitions, fait toutefois apparaître l'existence d'une décomposition du tout (la structure) en parties et de relations entre ces parties. Les méthodes structurelles de reconnaissance de formes adoptent des représentations faisant explicitement apparaître les relations entre éléments de la forme . Elles sont évidemment plus

appropriées pour traiter des structures visuelles où les composantes sont liées par des relations spatiales.

La décomposition d'une forme peut faire apparaître une hiérarchie de composantes. On appelle primitives les composantes élémentaires, l'information de granularité la plus fine qui ne peut plus être décomposée (le niveau d'information de ces primitives est fixé par le concepteur du système de reconnaissance).

Une première phase de la reconnaissance sera donc de ramener la donnée brute à un ensemble de composantes, c'est-à-dire de la segmenter en fragments, suivant un ou plusieurs niveaux de décomposition. Les traits descriptifs d'une structure sont de type **fragment** mais aussi des traits qui expriment des **configurations** comme la symétrie ou la fermeture, ou des **événements** qui se produisent sur la forme comme un changement de direction au niveau d'un point anguleux ou encore des **caractéristiques** comme le rapport hauteur sur largeur, les moments, le centre de gravité, ... toute information qui caractérise la structure globale.

Les **relations** entre les composantes peuvent être explicitées par différents formalismes de représentation. Le plus simple est la chaîne de composantes, particulièrement adaptée au cas des signaux monodimensionnels et des contours. Pour des relations spatiales dans l'espace 2D ou 3D, des représentations sous formes d'arbres ou plus généralement de graphes sont définies.

a) Comparaison dynamique de chaînes de symboles

La comparaison de deux formes représentées par des chaînes de symboles se ramène à calculer des distances entre chaînes. Sachant que les deux chaînes ne sont pas identiques, il faudra définir des opérations et des coûts associées pour transformer ces chaînes afin de les rendre identiques et pour évaluer numériquement le coût de cette transformation.

Les opérations d'édition :

Opérations	Coûts
$a \rightarrow b$ Substitution de a par b	$\gamma(a,b)$
$a \rightarrow \lambda$ Elimination de a	$\gamma(a,\lambda)$
$\lambda \rightarrow a$ Insertion de a	$\gamma(\lambda,a)$

Exemple :

$$X = a b \quad ; \quad Y = b a c$$

Opérations S :

$$\lambda \rightarrow c \quad b \rightarrow \lambda \quad c \rightarrow b \quad \lambda \rightarrow c$$

$$ab \Rightarrow cab \Rightarrow ca \Rightarrow ba \Rightarrow bac$$

Coût :

$$\gamma(S) = \gamma(\lambda, c) + \gamma(b, \lambda) + \gamma(c, b) + \gamma(\lambda, c)$$

L'algorithme de Wagner et Fisher est un algorithme de programmation dynamique qui calcule la distance d'édition minimale entre deux séquences de symboles. Il peut être généralisé à des séquences de symboles attribuées ou de vecteurs. D'autres algorithmes de ce type existent, en particulier en parole pour comparer des mots. Il existe aussi des variantes de W et F pour rechercher des motifs (des sous-séquences) dans des séquences, pour comparer des suites de symboles de manière cyclique, des polygones indépendamment de leur orientation etc.

La comparaison de graphes est comparable par beaucoup d'aspects à la comparaison de chaînes. Un des problèmes est le nombre de possibilités qui existent pour mettre en correspondance deux graphes (l'ordre des symboles imposé par les représentations séquentielles limite le nombre de comparaisons de symboles, ce qui n'est pas le cas avec les graphes en général). On accélère le temps de calcul et on réduit l'espace mémoire où sont stockées les comparaisons partielles en utilisant des heuristiques (algorithme A* par exemple).

b) Modèles de Markov

Les modèles de Markov généralisent la comparaison dynamique de séquences en introduisant un modèle stochastique de la classe. Assigner une séquence à une classe se ramène alors à calculer la probabilité qu'elle ait été engendrée par les modèles des classes.

La séquence est considérée comme un signal observable dans le temps, un état est associé à chaque instant d'observation. Le modèle de Markov est décrit par l'ensemble de ses états, les probabilités de transitions entre états et la probabilité d'observation par état.

Pour un modèle à N états d'observations, on note $s_i, i=1 \dots N$ les états et $q_t, t=1 \dots T$ la variable aléatoire qui prend ses valeurs dans l'ensemble des observations. Un processus stochastique vérifie la propriété de Markov d'ordre 1 si : pour tout t,

$$P(q_t=s_i / q_{t-1}=s_j, q_{t-2}=s_k \dots) = P(q_t=s_i/q_{t-1}=s_j)$$

L'évolution du processus est entièrement déterminée par la probabilité initiale et les probabilités de transition entre états.

c) Modèle à états cachés :

Dans ce cas, on a un double processus stochastique. Les états ne correspondent plus directement à des observations. Pour chaque état caché q_i , il y a une probabilité d'observer une réalisation v_k parmi $\{v_1, v_2, \dots, v_M\}$.

Problèmes :

- Evaluer la probabilité d'observation d'une séquence, observée soit $P(O/I)$ pour une suite d'observation O et un HMM $I=(A, B, p)$ où A représente les probabilités de transitions, B celles des observations et p celles de l'état initial.
- Recherche du chemin le plus probable, ou décision (algorithme de Viterbi).
- Apprentissage : ajuster les paramètres de I pour maximiser $P(O/I)$. C'est le problème clé, plusieurs algorithmes dont EM (Estimation-Modification) qui permet d'apprendre en deux temps les paramètres I et la segmentation en états.

Applications : Parole, écriture, gestes 3D.



Chapitre II

L'Image sur ordinateur

1. Introduction

La vision nous permet de percevoir et d'interpréter le monde qui nous entoure. La vision artificielle a pour but de reproduire certaines fonctionnalités de la vision humaine au travers de l'analyse d'images. C'est un problème difficile en raison du fait que la machine ne voit dans l'image qu'un ensemble de points.

Dans la suite de ce chapitre on tentera de présenter l'image sur machine autrement dit on détaillera le concept de l'image numérique.

2. L'image numérique

2.1. L'image sur ordinateur

Une image numérique est ensemble de pixels contenant chacun différentes informations (intensité lumineuse, couleur...). Ces informations sont codées dans une grille échelonnée. Selon la nature de cette grille on peut obtenir différents types d'images.

2.2. Les type d'images

2.2.1. Image noir et blanc (monochromes)

Ces images sont dites à niveau de gris, car on ne prend pas en compte la couleur mais seulement l'intensité lumineuse (l'exemple classique correspond aux photographies noir et blanc). Parmi ces images on peut trouver :

- **Image binaire** : Dans l'image binaire chaque pixel est représenté par un bit (0/1) avec en général (0 = noir, intensité nulle et 1= blanc, intensité maximale). Notons que la plupart des systèmes de traitement d'images placeront chaque pixel dans un **octet** pour des facilités d'accès et d'écriture des algorithmes.



Figure.2 : Image d'un visage en noir et blanc (binaire)

- **Images en niveaux de gris** : Dans ce type d'image on dispose d'une échelle de teintes de gris, et la plupart du temps on dispose de 256 niveaux de gris avec :

0 = Noir,.....,127 = gris

128 = Moyen,.....,255 = Blanc



Figure. 3 : Image d'un visage en niveaux de gris

2.2.2. Images couleur

Il existe plusieurs approches qui débouchent sur la création de nombreux espaces de couleurs. On peut principalement distinguer l'approche additive et l'approche soustractive, qui sont parfois imposés par le support sur lequel on veut visualiser l'image.

Le modèle additif est fondé sur 3 couleurs dites primaires telles que toutes les autres couleurs que l'on veut créer peut être obtenue par addition de quantités plus ou moins grandes de ces 3 couleurs. Le modèle additif connu est celui dont les 3 couleurs primaires sont le rouge, le verte, et le bleu ou modèle RVB (en anglais RGB).

Le modèle soustractif à comme couleurs primaires, le Cyan, le magenta et le jaune, on l'appelle aussi modèle CMY. [1]

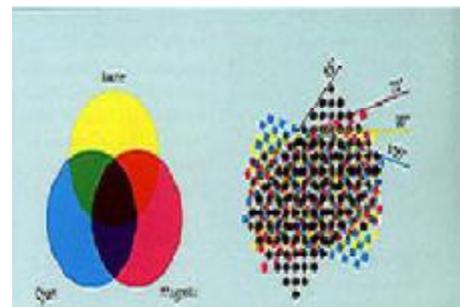
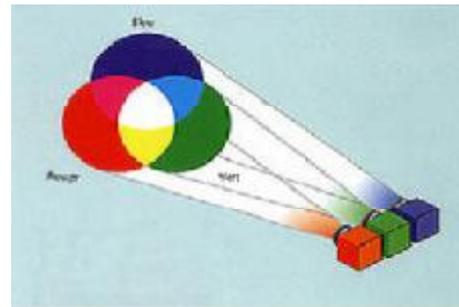


Figure. 4 : Modèle additive et soustractive des couleurs



Figure. 5 : Image couleur d'un visage

3. Représentation de l'image numérique

L'image est manipulée par l'ordinateur sous forme d'un fichier composé généralement de deux parties :

- § L'entête qui contient des informations générales sur l'image (hauteur, longueur, nombre de couleurs utilisées, ... etc.). Ce n'est pas le cas pour tous les types d'images, l'entête des fichiers TIFF contient un index vers une zone contenant des informations particulières.
- § L'image proprement dite qui est une structure de données contenant les valeurs de chaque pixel (généralement une matrice).

Parmi les formats standards de fichier image on peut citer :

Format	Compression	Dimensions maximales (pixel)	Nombre de couleurs maximales
BMP	aucune / RLE	65 536 x 65 536	16 777 216
GIF	LZW	65 536 x 65 536	256
IFF	aucune / RLE	65 536 x 65 536	supérieur à 16 777 216
JPEG	JPEG	65 536 x 65 536	supérieur à 16 777 216
PCX	aucune / RLE	65 536 x 65 536	16 777 216
PNG	RLE	65 536 x 65 536	supérieur à 16 777 216
TGA	aucune / RLE	65 536 x 65 536	supérieur à 16 777 216
TIFF/TIF	Packbits / CCITT G3&4 / RLE / JPEG / LZW / UIT-T	$2^{32}-1$	supérieur à 16 777 216

Tableau 1 Les différents formats de stockage des images

4. Caractéristiques d'une image numérique

4.1. Résolution d'image

On a vu qu'une image numérique peut être représentée par une matrice de valeurs numériques avec **N** lignes et **P** colonnes. Le couple constitué par le nombre de pixels (un pixel est un point élémentaire d'une image, sa valeur reflète soit un niveau de gris ou une couleur) **NxP** est appelé résolution de l'image.

4.2. Image à niveaux de gris

C'est une image où est utilisée **N+1** couleurs obtenues après dégradation de la couleur noire. Après la quantification, le nombre de niveaux est déterminé (le niveau 0 pour le noir et le niveau **N** pour le blanc).

Pour $N=1$ (deux niveaux 0 pour le noir et 1 pour le blanc) l'image est appelée "monochrome".

4.3. Image en couleurs

C'est une image où est utilisée les trois couleurs standards : le rouge (Red), le vert (Green) et le bleu (Blue) qui sont connues sous le terme RGB. En faisant combiner ces trois couleurs, on peut générer d'autres couleurs qui peuvent être codées sur 1 octet (256 couleurs), sur 2 octets (65536 couleurs) ou sur 3 octets (2^{24} couleurs).

La valeur qui reflète l'intensité lumineuse d'un pixel est alors vue comme étant un vecteur tridimensionnel dans l'espace formé par les trois couleurs essentielles.

4.4. Bruit :

Le bruit est considéré comme un phénomène de brusque variation d'un pixel par rapport à ses voisins suivant un certain seuil (la différence moyenne entre la valeur du pixel en question est celles de ces voisins dépasse le seuil). Il peut provenir de l'éclairage, des dispositifs optiques ou électroniques du capteur ou autres.

4.5. Effets de bords : [1]

Ce phénomène concerne les pixels se trouvant aux frontières de l'image. En effet, si on applique une méthode qui traite chaque pixel de l'image en fonction de ses voisins, on remarque que ce n'est pas possible aux pixels qui constituent le cadre extérieur :

Les solutions utilisées pour résoudre ce problème sont alors de :

- § Traiter uniquement la partie intérieure de l'image, mais cela fait diminuer la taille de cette dernière de la taille du masque utilisé -1 (si la résolution de l'image initiale est de $n \times p$, et si la taille du masque est de $s \times s$ alors la taille de l'image résultat est $[n-s+1] \times [p-s+1]$).
- § Réaliser un effet dit de miroir, il consiste à ajouter autour de l'image initiale un cadre d'épaisseur déterminée selon le masque utilisé (taille du masque/2) dans lequel on recopie les lignes de bords de l'image ou bien le remplir par des zéros.

4.6. Luminance : [3]

La luminance est l'intensité lumineuse par unité de surface perpendiculaire à la direction d'émission (émission du flux lumineux).



4.7. Contraste [1]

Si E_1 et E_2 sont les valeurs des éclairagements de deux plages voisines d'une image, le contraste C est par définition :

$$C = \frac{E_1 - E_2}{E_1 + E_2}$$

Une image est dite contrastée si les variations entre les pixels voisins sont faibles, ainsi les détails de l'image auront peu de chance d'être visibles car la transition entre régions s'étale sur plusieurs pixels.

4.8. Contours et textures [1]

Une image numérique a été vue comme étant une suite de points dont chacun possède une valeur numérique. Une manière plus moderne et plus naturelle de représenter une image est de la modéliser en termes de contours et de textures. Les contours représentent la frontière entre les objets de l'image alors que les textures décrivent la structure de ceux-ci.

5. Traitement et Analyse d'image

Le processus classique de découpage de l'analyse d'image est comme suit :

Ø **Acquisition** : Acquisition de l'image grâce à un appareillage approprié.

Scène physique \mathcal{P} représentation numérique (image numérique).

Ø **Filtrage** : amélioration de l'image.

Image numérique bruitée \mathcal{P} Image numérique plus nette

Ø **Segmentation** : Extraction de l'information pertinente par segmentation

Image numérique \mathcal{P} description structurelle de l'image.

Ø **Interprétation** : donner une interprétation à la scène.

Description structurelle \mathcal{P} description sémantique.

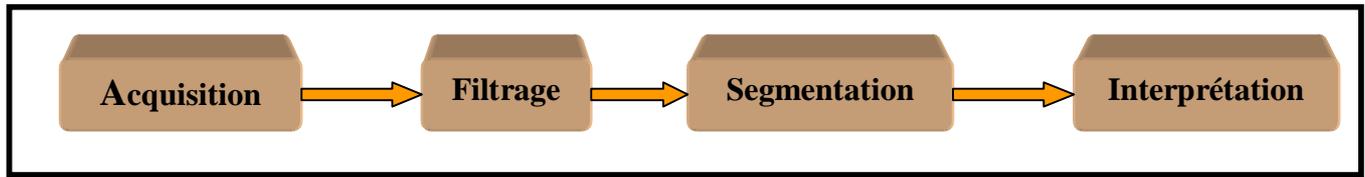


Figure.6 : Processus de traitement d'image

5.1. Acquisition

La mise en œuvre d'une expérimentation nécessite le plus souvent la production d'images numériques adaptées au thème qu'on souhaite étudier. L'acquisition est donc la première étape à surmonter.

Les deux caractéristiques principales des images, lors de l'acquisition, sont l'échantillonnage et la quantification :

- ✓ L'*échantillonnage* spatial détermine la taille de chaque point élémentaire de l'image (pixel). Cette taille est fonction de la résolution du capteur. Les contraintes physiques imposent ensuite les dimensions maximales de l'image (nombre de pixels en ligne et en colonne).
- ✓ La *quantification* est le résultat de la conversion de la mesure en une valeur discrète (entière). Elle fixe le nombre de niveaux possibles pour chaque pixel de l'image, niveaux souvent interprétés comme des niveaux de gris.

L'acquisition et le stockage d'images sur ordinateur est aujourd'hui un procédé banal. De nombreux outils à faible coût les rendent accessibles sur les ordinateurs personnels. Parmi ces outils on cite :

Ø *Numérisation des documents analogiques* : Un scanner est un dispositif à balayage qui analyse point par point la lumière réfléchie par un document. A partir d'un document photographique en couleurs, la numérisation avec un scanner produit en général un triplet de valeurs numériques pour chaque pixel, triplet dont la signification dépend du modèle de représentation de la lumière choisi.

Ø *Photographie numérique* : La photographie numérique intègre dans un même appareil le dispositif optique de prise de vue et l'électronique de numérisation. Le

Canon Ion stocke sur une disquette d'un format spécial une image vidéo analogique qui peut être ensuite numérisée avec une carte de conversion vidéo-numérique. Kodak propose un service de conversion des images photographiques en format numérique propriétaire sur support CD-ROM ou micro-disquette... Les appareils entièrement numériques sont basés sur la technologie CCD que l'on retrouve dans les caméras vidéo.

Ø **Numérisation d'images vidéo** : Les cartes de numérisation vidéo offrent une grande variété de possibilités qui sont liées à l'origine du signal vidéo et au standard utilisé (PAL, SECAM, NTSC). La qualité de l'image dépend en particulier du type de signal d'entrée analogique. La résolution des images vidéo dépend du nombre de lignes permettant le balayage d'une image complète. Dans le domaine du multimédia, de gros progrès ont été réalisés récemment pour fournir des séquences vidéo numérisées très compactes.

L'acquisition d'images numériques mosaïques (sous forme de matrices de nombres) transforme un signal continu analogique (une radiation lumineuse) en échantillonnage numérique discontinu. Il y a donc des pertes d'information dues aux méthodes d'échantillonnage et d'autres dues aux erreurs d'acquisition (capteur défectueux, couverture nuageuse, défauts d'optique, de parallaxe, etc.). D'où les traitements et les filtrages d'images.

5.2. Filtrage

Le filtrage d'images qui constitue une étape de prétraitement regroupe l'ensemble des processus visant à améliorer les caractéristiques d'une image. En d'autres termes on cherche à minimiser l'influence du bruit dans l'image, l'origine de bruit est produit à cause de plusieurs facteurs par exemple le contexte d'acquisition (Sur ou sous illumination, Perturbations des capteurs), ou la nature du capteur (Distorsions : géométriques, d'intensité). En ce sens on cherche donc d'une certaine manière à moyennner localement l'image [2].

Il existe différents types de filtrage les plus connus sont :

a. Filtre linéaires

§ **Le filtre moyen** : C'est l'opérateur de lissage la plus simple et le plus connu, il associe à un pixel de l'image la moyenne de ses huit pixels voisins, donc une

combinaison linéaire, normalisé par les facteurs 1/9 ou 1/16 pour ramener les valeurs entre 0 et 255, deux exemples de filtre de lissage (filtre passe bas) pour un masque de matrice de convolution 3*3 peuvent être donnés :

$$\frac{1}{9} * \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad \text{OU} \quad \frac{1}{16} * \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

Pour le premier masque, l'image filtrée aura pour expression :

$$I_F = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 H_1(i+1, j+1) * I_I(x+i, y+j) [3].$$

b. Filtre non linéaire

Contrairement au type précédent de filtrage, la valeur du pixel n'est pas la combinaison linéaire des pixels voisins, les lissages linéaire des ont pour inconvénient de traiter de la même façon les signaux utiles, les bruits et les parasites qui sont superposés.

L'utilisation des algorithmes de lissage non linéaire permet de supprimer ou d'atténuer les valeurs trop différentes de celles des pixels voisins [4].

§ Le filtrage médian: La fenêtre des pixels formant le voisinage contenant un nombre impair de pixels, on donne au pixel central la valeur médiane de l'ensemble des pixels de cette fenêtre.

Le principe consiste à faire le tri par ordre croissant des neufs pixels voisins et d'affecter la valeur médiane au pixel considéré (voir figure 7) [3].

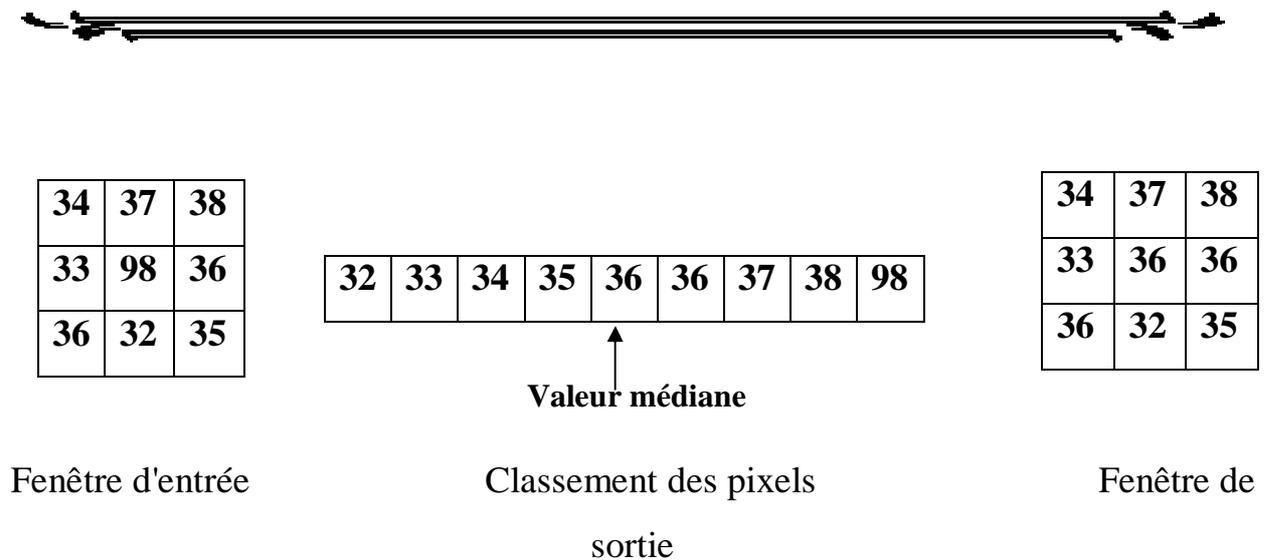


Figure 7 : Applications d'un filtre médian 3*3 sur une fenêtre 3*3

§ **Filtre minimum et maximum :** pour les filtres minimum et maximum, on applique le même principe que celui du filtre médian mais le pixel central dans ce cas sera remplacé par la valeur minimale ou la valeur maximale, selon le filtre, de l'ensemble des pixels de la fenêtre de voisinage [3].

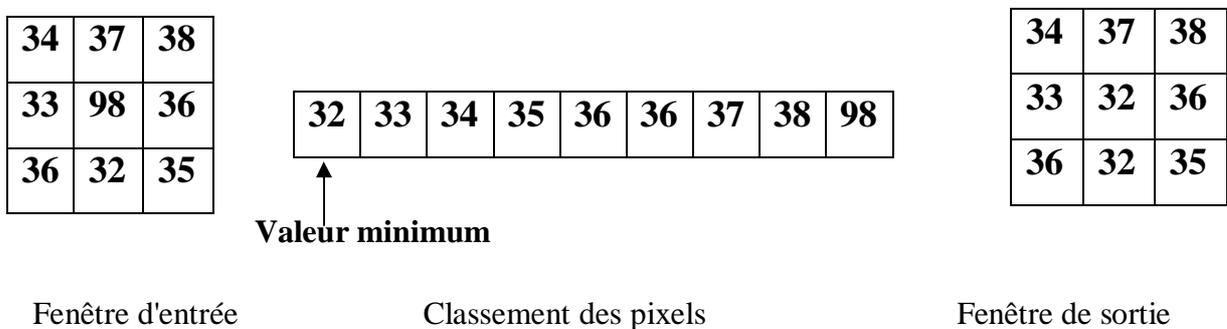


Figure. 8 : Applications d'un filtre minimum sur une fenêtre 3*3

Comme indiqué précédemment, ce type de filtre utilise une loi non linéaire du voisinage d'un pixel pour le modifier. Ceci est réalisé en faisant balayer tous les pixels de l'image ou une grande partie d'entre eux par masque dont les coefficients dépendent du filtre utilisé. Le pixel central du masque est celui qui est affecté par la transformation.



5.3 Segmentation d'Image :

La segmentation est une technique de traitement d'image qui consiste à extraire de l'image des indices visuels soit des contours soit des régions, on trouve plusieurs approches de segmentation qui peuvent être classer en deux grandes catégories: La segmentation par détection de contours et la segmentation par détection de régions.

En traitement d'image, la segmentation peut être définie comme le processus de découpage d'une image en zones homogènes séparer par des bords et traversées par des failles, pour construire une série des régions telles que :

$$R = \bigcup_{i=0}^{N-1} R_i \quad R_i \cap R_j = \emptyset, i \neq j$$

Et on peut abordée la segmentation sous l'angle local ou globale, la premier (local) considère que la segmentation ne requiert pas la connaissance de contenu globale de l'image, mais seulement le voisinage directe de chaque pixel. A l'inverse de l'approche globale qui cherche la totalité de contenu de l'image [8].

5.3.1 Segmentation par extraction de régions :

La segmentation des images par détection des régions consiste à regrouper les pixels de ces images qui partagent une même propriété pour former des régions connexes. Cette section sera détaillée dans le prochain chapitre.

5.3.2 Segmentation par extraction de contours :

Un contour est une zone de transition séparant deux régions homogènes présentant des caractéristiques différentes. Un contour local est une petite surface de l'image dans laquelle les niveaux de gris changent rapidement dans une direction donnée.

Un opérateur de contour est un opérateur mathématique développé pour détecter la présence de discontinuités locales dans la fonction de luminance d'une image [5]. Il existe plusieurs méthodes de détection de contours les plus connues sont :

a. Méthodes différentielles

L'application des opérateurs de différentiation permet une accentuation des hautes fréquences, très riches en information sur les contours. L'extraction de contour s'effectue classiquement de deux manières : l'approche du Gradient, et l'approche du Laplacien.

- **Approche du gradient**

Le principe de base de détection de contours dans une direction donnée est le calcul de la différence d'intensités de pixels adjacents dans la direction orthogonale à celle choisie. Pour une détection de contours de sens horizontale, on effectue la différence suivante: $I(x,y) - I(x+1,y)$

Une valeur non nulle de cette différence indique qu'il y a présence d'un contour local. Le gradient est généralement exprimé sous forme de produit de convolution de l'image par un masque. Les coefficients de ces masques peuvent être ajustés pratiquement selon les besoins. C'est d'ailleurs de façon totalement empirique que la plupart des masques courants ont été introduits [8].

Les opérateurs les plus connus sont :

Opérateur de	Approximation	Masque
Roberts	Roberts approxime les dérivées directionnelles suivant les axes orientées à 45 degrés à l'aide d'une convolution de l'image $f(x,y)$ avec des masques 2*2	$H_0 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$ $H_1 = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix}$
Prewitt et Sobel	Prewitt et Sobel approximent les dérivées directionnelles suivant les axes O_x et O_y à l'aide d'une convolution de l'image $f(x,y)$ avec les masques 3*3 suivant:	$H_x = \begin{vmatrix} 1 & 0 & -1 \\ c & 0 & -c \\ 1 & 0 & -1 \end{vmatrix}$ $H_y = \begin{vmatrix} -1 & -c & -1 \\ 0 & 0 & 0 \\ 1 & c & 1 \end{vmatrix}$

- **Approche du Laplacien**

Le Laplacien est une approximation spatiale de la dérivée seconde, il a l'avantage d'être invariant par rotation, des approximations de ce filtre sur une fenêtre de 3 * 3 sont définies par les masques suivants [7] :

$$H1 = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix} ; H2 = \begin{vmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{vmatrix} ; H3 = \begin{vmatrix} -1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{vmatrix}$$

b. Méthode morphologique :

Ce filtre est un filtre morphologique qui consiste à faire une succession d'alternances d'ouvertures et de fermetures par des sphères de rayons croissants.

Filtre (f)= ouverture (fermeture (f)) [1].

5.4. Interprétation

Le traitement d'une image est souvent envisagée comme une procédure de décomposition en régions homogènes qu'il s'agit ensuite d'identifier ou de reconnaître. A cette fin, il est souvent impérative de dégager un ensemble de caractéristiques qui nous permettent de donner une interprétation à l'image.

6. Domaine d'application

Une caractéristique très importante et en même temps très enrichissante du traitement d'image est son aspect interdisciplinaire. On trouve ses applications dans des domaines très variés tels que :

- les télécommunications (télévision, vidéo, transmission, stockage, téléconférence, services intégrés RNIS, publicité),
- la médecine, (radiographie, tomographie, ultrasons, microscopie, fichier électronique et diagnostics à distance),
- l'industrie (robotique, surveillance de qualité, sécurité),



- la biométrie,
- ...

La liste est longue et s'allonge chaque jour davantage.

Chapitre III

La biométrie et la reconnaissance des visages humains

1. Définition

La biométrie est un ensemble des technologies (appelées les technologies biométriques) qui exploitent des caractéristiques humaines physiques ou comportementales telles que l'empreinte digitale, la signature, l'iris, la voix, le visage, la démarche, et un geste de main pour différencier des personnes. Ces caractéristiques sont traitées par certain ordre des processus automatisés à l'aide des dispositifs comme des modules de balayage ou des appareils-photo. Á la différence des mots de passe ou des PINs (numéros d'identification personnelle) qui sont facilement oubliés ou exposés à l'utilisation frauduleuse, ou des clefs ou des cartes magnétiques qui doivent être portées par l'individu et sont faciles à être volées, copiées ou perdues, ces caractéristiques biométriques sont uniques à l'individu et il y a peu de possibilité que d'autres individus peuvent remplacer ces caractéristiques, Donc les technologies biométriques sont considérées les plus puissantes en termes de sécurité.

En plus les mesures biométriques sont confortables parce qu'elles n'ont pas besoin d'être portées séparément. De telles caractéristiques peuvent être bien employées pour obtenir l'identification/authentification pour accéder à des systèmes tels que les ATMs (guichet automatique). La biométrie se prouve également comme outil puissant d'identification/vérification aux scènes de crime dans le secteur juridique [10].

2. Système biométrique

Un système biométrique est essentiellement un système de reconnaissance de formes qui fonctionne en acquérant des données biométriques à partir d'un individuel, extrayant un ensemble de caractéristiques à partir des données acquises, et comparant ces caractéristiques

contre la signature dans la base de données. Selon le contexte d'application, un système biométrique peut fonctionner en mode de vérification ou mode d'identification:

- **Vérification** : le système valide l'identité d'une personne en comparant les données biométriques capturées à sa propre base de données. Dans un tel système, un individu qui désire être identifié réclame une identité, habituellement par l'intermédiaire d'un PIN (numéro d'identification personnelle), d'un nom d'utilisateur, d'une carte fûtée, etc., et le système Biométrie pour l'identification conduit une comparaison d'un-à-un pour déterminer si la réclamation est vraie ou faux (Est-ce que ces données biométriques appartiennent à Bob?).
- **Identification**: le système identifie un individu en recherchant les signatures (Template) de tous les utilisateurs dans la base de données. Par conséquent, le système conduit plusieurs comparaisons pour établir l'identité d'un individu (ou échoue si le sujet n'est pas inscrit dans la base de données de système) [10].

3. Erreurs du système biométrique

Un système biométrique fait deux types d'erreurs:

(1) confusion des caractéristiques biométriques de deux personnes différentes et les identifier comme la même personne.

(2) confusion des caractéristiques biométriques de la même personne pour être de deux personnes différentes. Ces deux types d'erreurs se nomment souvent en tant que faute d'acceptation et faute de rejet, respectivement [12].

4. Applications de la biométrie

Les techniques biométriques sont appliquées dans plusieurs domaines et son applications sont divisées en trois groupes principaux :

- Û Applications commerciales: telles que l'accès au réseau informatique, la sécurité de données électroniques, le commerce électronique, l'accès d'Internet, l'ATM, la carte de crédit, le contrôle d'accès physique, le téléphone portable, le PDA, la gestion des registres médicale, l'étude de distance, etc...
- Û Applications de gouvernement: telles que la carte nationale d'identification, le permis du conducteur, la sécurité sociale, la contrôle de passeport, etc...
- Û Applications juridique: telles que l'identification de cadavre, la recherche criminelle, l'identification de terroriste, les enfants disparus, etc...

5. Techniques biométriques

En fait Il y a trois catégories de technique de biométrie:

- a. Les techniques d'analyse du comportement: il s'agit d'un type de biométrie caractérisées par un trait d'attitude qui est appris et acquis au fil du temps, Par exemple la dynamique de la signature, La façon d'utiliser un clavier d'ordinateur, etc...).
- b. Les techniques d'analyse de la morphologie humaine: il s'agit d'un type de biométrie défini par les caractéristiques physiques (empreintes digitales, forme de la main, forme du visage, dessin du réseau veineux de l'œil, la voix, etc...) [11].
- c. L'analyse de trace biologique (odeur, salive, ADN, urine....)[13]

En résumé, **la biométrie constitue un identificateur de personne qui ne peut être volé, oublié ou dupliqué.**

L'unité biométrique sur laquelle se basera le système d'identification doit remplir certaines conditions :

- 1- **L'unicité** : il est nécessaire qu'elle soit unique pour chaque personne, deux individus ne peuvent pas avoir exactement la même voix ou ADN
- 2- **Universelle** : toute la population doit posséder cette biométrie.
- 3- **Mesurable** : le système doit pouvoir la quantifier.
- 4- **Enregistrable** : peut être emmagasinée.

6. L'industrie de la biométrie

Le marché des produits d'authentification individuelle par l'approche biométrique est en forte croissance, certe la technologie actuellement dominante est celle employant les empreintes digitales, la raison en est simple : on accepte depuis longtemps le fait que les empreintes digitales soient uniques pour chacun d'entre nous [14], néanmoins, après les empreintes digitales, la tendance se tend de plus en plus vers le visage, comme l'illustre la figure 9.

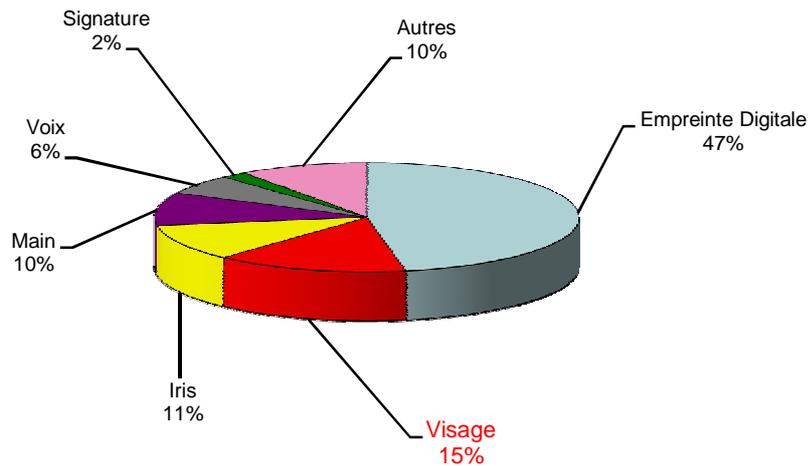


Figure 9 : Marché des systèmes d'identification

7. Quelques techniques biométriques

I. La reconnaissance d'iris

La reconnaissance de l'iris est considérée comme la technique biométrique la plus exacte. L'iris est la partie colorée de l'œil qui entoure la pupille noire. Son inspection attentive révèle de nombreuses structures détaillées. Ces structures se développent dès le plus jeune âge et sont considérées comme uniques, même chez les vrais jumeaux. Les structures de l'iris resteront constantes jusqu'à la mort.

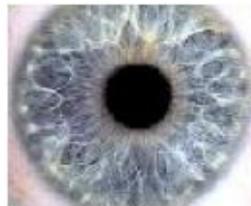


Figure 10: Image d'iris

Pour reconnaître un individu, on extrait des paramètres de l'image de l'œil. Puis, on compare ces paramètres avec tous les paramètres précédemment extraits et sauvegardés. La comparaison qui produira le meilleur score permettra d'identifier le nom de l'individu.

Ü Utilisations actuelles:

La reconnaissance de l'iris est utilisée dans le secteur financier pour les employés et les clients, pour le téléchargement de musique par Internet, dans les guichets automatiques (ATM Bank United), pour le paiement dans les supermarchés (Kroger du Texas), dans les institutions carcérales, dans les hôpitaux et dans les aéroports [62].

II. Empreintes digitales

Empreinte digitale est la caractéristique d'un doigt. On le croit que chaque empreinte digitale est unique. Chaque personne a ses propres empreintes digitales avec l'unicité permanente. Ainsi les empreintes digitales sont utilisées depuis longtemps pour l'identification et l'investigation juridique.

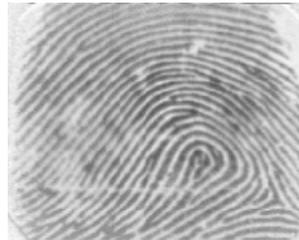


Figure 11: Empreinte digitale acquise par un capteur optique

Une empreinte digitale se compose de beaucoup des rides et sillons. Ces rides et sillons présentent de bonnes similitudes dans chaque petite fenêtre locale, comme le parallélisme et la largeur moyenne. Cependant, montré par des recherches sur l'identification d'empreinte digitale, des empreintes digitales ne sont pas bien distinguées par leurs rides et sillons, mais par des minuties qui sont des points anormaux sur les rides.

Biométrie pour l'identification Parmi les types des minuties rapportés dans la littérature, deux types sont les plus utilisés ce sont l'arrêt de ride, la fin d'une ride, et la bifurcation, le point sur la ride de laquelle deux branches dérivent (figure.12)



Figure 12 : Deux types de minuties les plus utilisés dans la littérature

Ü Utilisations actuelles :

Ce type de système est utilisé par les institutions financières pour leurs employés et leurs clients. Il se retrouve également dans les magasins de détail, les hôpitaux, les écoles, les aéroports, les cartes d'identité, les passeports, les permis de conduire et de nombreuses autres applications [62].

III. Signature

Signer un document pour s'identifier est un geste naturel pour tout le monde. Dans la vie de tous les jours, nous signons régulièrement des documents. La reconnaissance de la signature est une des techniques biométriques comportementales. Dans le système d'identification par la signature, l'utilisateur doit signer avec un stylo lecteur sur une tablette graphique, le système analyse ensuite les variations de vitesse du stylo, la différence de pression sur la tablette graphique au cours de la signature. Le point faible de cette technique est qu'un individu qui ne signe pas toujours de la même façon se verra souvent refuser l'accès au système. L'acceptation de cette technique est très bonne car la signature est un geste commun pour tout le monde.

Utilisations actuelles :

Ces systèmes sont utilisés dans les compagnies pharmaceutiques, les prisons, les services postaux et les banques [62].

IV. La reconnaissance de Visage

La reconnaissance de la forme du visage est la technique de Biométrie pour l'identification la plus commune et la plus populaire. Elle est aussi la plus acceptable parce que on peut l'utiliser à distance sans la collaboration avec d'autre objet.

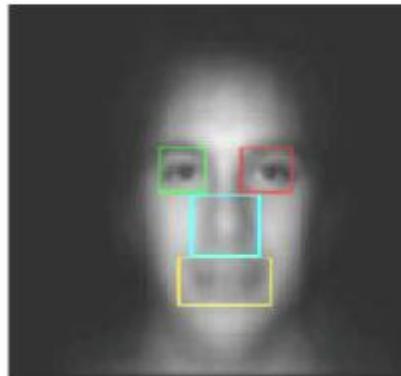


Figure 13: Reconnaissance de visage

Utiliser une caméra permet d'acquies la forme du visage d'un individu et puis retirer certaines caractéristiques. Les caractéristiques essentielles pour la reconnaissance du visage sont: les yeux, la bouche, le tour du visage, le bout du nez, etc. Selon le système utilisé, l'individu doit être positionné devant la caméra où peut être en mouvement à une certaine distance. Les données biométriques qui sont obtenues sont comparées au fichier référence.

Le logiciel doit être capable d'identifier un individu malgré différents artifices physiques (moustache, barbe, lunettes, etc..).

Le visage est une biométrie relativement peu sûre. En effet, le signal acquis est sujet à des variations beaucoup plus élevées que d'autres caractéristiques. Celles-ci peuvent être causées, entre autres, par le maquillage, la présence ou l'absence de lunettes, le vieillissement et l'expression d'une émotion. La méthode de la reconnaissance du visage est sensible à la variation de l'éclairage et le changement de la position du visage lors de l'acquisition de l'image. Dans le *monde physique*, il y a trois paramètres à considérer à l'acquisition de l'image : L'éclairage, la variation de posture et l'échelle. La variation de l'un de ces trois paramètres peut conduire à une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents. Le *codage* consiste en l'acquisition d'image et sa digitalisation, il comporte un risque de bruit et donne lieu à une représentation 2-D (la matrice des niveaux de gris) pour un objet 3-D (le visage) [11].

La reconnaissance du visage est utilisée comme système de surveillance ou d'identification par les autorités ou les corps policiers principalement dans les lieux publics, les aéroports, les frontières, les casinos, les plages, les guichets automatiques et les laboratoires [62].

IV.1. Caractéristiques visage

IV.1.1 Définition et particularités du visage :

Le visage peut être défini comme :

« Une structure tridimensionnelle possédant une configuration "externe" – le contour de la face modelé par les saillies osseuses et souligné par les cheveux – où s'inscrit une configuration "interne" formée par l'assemblage des traits. Certaines particularités s'y ajoutent comme la pilosité, les lunettes, la texture, la coloration de la peau... » [15].

Mais, sauf exception, tous les visages ont une structure voisine ne différant que sur d'infimes modifications de la configuration externe ainsi que de la forme des éléments internes et de leurs rapports. Et pourtant, malgré ces faibles différences, on reconnaît des milliers de visages, d'autant plus qu'un visage donné ne se présente jamais de la même façon...

Donc, à la base, la constitution du visage vient du regroupement des attributs faciaux, c'est-à-dire des parties distinctes et indépendantes d'un visage, pouvant être nommées et reconnues (yeux, nez, bouche....) , bien que les mêmes attributs se retrouvent dans chaque visage, les caractéristiques des ces parties constituantes peuvent parfois à elles seules suffir à la reconnaissance du visage , par exemple la forme du nez, la couleur des yeux ; le style de coiffure, on parle alors de reconnaissance par attributs , effectuée par traitement analytique du visage , cependant, les études en reconnaissance de visages ont rapidement démontré que le traitement analytique semble être accompagné d'un autre type de traitement potentiellement plus efficace : ce deuxième type de traitement reposerait sur l'information sous-jacente aux simples attributs faciaux, soit l'information configurale ou de deuxième ordre, dont l'importance relative varie selon les auteurs [16] ; De prime abord, la définition de ce qui constitue l'information configurale n'est pas nette et, mis à part le fait qu'elle repose sur les relations spatiales entre les différents attributs, plusieurs définitions peuvent être utilisées.

L'information configurale peut comprendre une très petite étendue du visage (p. ex., entre deux attributs adjacents) ou s'étendre sur une étendue beaucoup plus grande, englobant des attributs séparés par de larges distances [17]. Finalement, selon d'autres auteurs, l'information configurale vient du traitement holistique du visage, où, tel un gabarit, celui-ci est représenté en tant que tout indivisible, et non pas en terme de la somme des attributs [18].

Beaucoup de chercheurs ont tenté de déterminer qu'elles sont les traits du visage qui permettent de l'identifier c'est-à-dire les traits faciaux permettant de distinguer les individus , il a été avancé que les caractéristiques jugées significatives pour la reconnaissance du visage sont : les yeux, la bouche , le nez et le menton [19] c'est-à-dire que leur priorité décroît dans le processus d'identification de haut en bas, pour [20] , le nez ne présente pas de particularité pour la vue de face.

IV.2 Pourquoi le visage ?

Le visage est un stimulus dont l'importance et la particularité ont été fréquemment remarquées et continuent de l'être pour les raisons suivantes :

- ∅ il est tout d'abord un élément prépondérant de la vie sociale , il nous permet d'identifier une personne et ainsi retrouver un ensemble d'informations sémantiques et épisodiques qui lui sont propres [3], par ailleurs , il nous apporte un grand nombre d'informations que l'on connaisse ou non la personne , il permet de ce fait de

déterminer si on a affaire à un homme, femme, enfant , adulte ou une personne âgée , ou encore , à une personne d'origine africaine ou asiatique (l'origine ethnique), il est également un élément de la communication des états émotionnels et source d'inférences sur la personnalité ou les compétences à l'insu même de la personne .

- Ø Par la fréquence à laquelle on le rencontre dans l'environnement et par son contenu riche en information sociale de premier ordre, le visage humain constitue un stimulus visuel de classe à part, en effet, il est sans conteste, la partie du corps la plus riche en informations.

- Ø L'identification à partir du visage répond fortement aux exigences sécuritaires et défensives de plusieurs applications réelles telles que le contrôle d'accès aux sites dits sensibles (aéroports, bâtiments gouvernementaux , télésurveillance....).

- Ø Malgré la complexité de la mise en œuvre de systèmes de reconnaissance de visages humains , ils (les visages) restent les mieux adaptés aux environnements non contrôlés, où la seule information disponible est le visage , tout simplement parce que le visage ne nécessite pas une interaction active avec le système de reconnaissance , contrairement à d'autres systèmes d'identification basés sur d'autres traits du corps tels que presser sa main ou ses doigts sur un dispositif spécifique , ou bien se positionner bien droit devant un capteur (identification iris/rétine) où une coopération entre l'utilisateur et le système est primordiale pour le processus de reconnaissance , ceci , fait des systèmes de reconnaissance de visages humains des systèmes passifs plus robustes et plus performants .

IV .3. D'où vient la difficulté ?

La reconnaissance de visage est un vrai défi en recherches allant de la psychologie cognitive au neurosciences en passant par l'intelligence artificielle, ce challenge est du essentiellement au fait que la reconnaissance de visages humains demeure un problème extrêmement difficile à résoudre comme l'ont montré les évaluations conduites par NIST [\[21\]](#) car :

- Ø les visages de personnes différentes ont globalement la même forme ; ils comportent tous un nez, une bouche , deux yeux et un menton ...approximativement dans la même

configuration , alors que les images d'un même visage peuvent fortement varier du fait de conditions d'éclairage, changement de posture , l'échelle , l'expression faciale, également parce que le visage est une biométrie dynamique qui varie beaucoup ,donc elle est relativement peu sûre , en effet, le signal acquis est sujet à des variations beaucoup plus élevées que d'autres caractéristiques , celles-ci peuvent être causées , entre autre , par la maquillage , la pilosité , le vieillissement , les cicatrices ,l'expression d'une émotion , ou les interventions chirurgicales sur le cartilage crânien.

- ∅ Il ne semble pas y avoir une caractéristique tirée du visage qui soit invariante à tous les changements que peuvent subir l'image du visage.
- ∅ le problème d'identification des jumeaux, et celui de déceler s'il s'agit réellement d'un visage ou d'un masque en silicone
- ∅ Un autre problème émerge automatiquement qu'est la **représentation et la description du visage** : la taille de l'image de visage constitue des données de très grandes dimensions, ceci influence automatiquement sur la dimensionnalité et la complexité de la reconnaissance.
- ∅ Une autre difficulté apparaît suivant que l'acquisition des images des visages se fait dans un environnement contrôlé ou non, dans un environnement non contrôlé, une série de prétraitements sont souvent nécessaires avant la reconnaissance proprement dite, chose qui augmente la complexité du problème.

Il est important de signaler que dans le but de simplifier cette tâche de reconnaissance qui est assez complexe, la plus part des concepteurs supposent certaines conditions sur les images à traiter et l'environnement, citons par exemple :

- ∅ L'image comporte un visage, soit une vue de face, ou de profil où tous les traits communs doivent sûrement figurer dessus.
- ∅ Le visage est vertical et sans ou avec une légère inclinaison.
- ∅ L'arrière plan est neutre.
- ∅ Pas d'occlusion.
- ∅ La personne ne porte pas de cicatrices, de déguisements

IV.4. Architecture générale d'un système de reconnaissance de visages

Comme tout processus de reconnaissance des formes, celui de la reconnaissance automatique des visages se décompose en deux étapes :

- 1 - La détection de la forme du visage dans l'image et la normalisation de celle-ci.
- 2 - L'identification du visage.

De plus, tout processus automatique de reconnaissance des visages doit prendre en compte plusieurs facteurs qui contribuent à la complexité de sa tâche, car le visage varie continuellement sous l'influence de plusieurs facteurs.

On ce qui nous concerne, nous nous focaliserons sur l'identification proprement dite, la détection quant à elle constitue un autre domaine de recherches, la section 6 donnera un aperçu sur cette étape qui présente à elles seul un axe de recherche, la figure 14 illustre la démarche générale adoptée pour réaliser de tels systèmes.

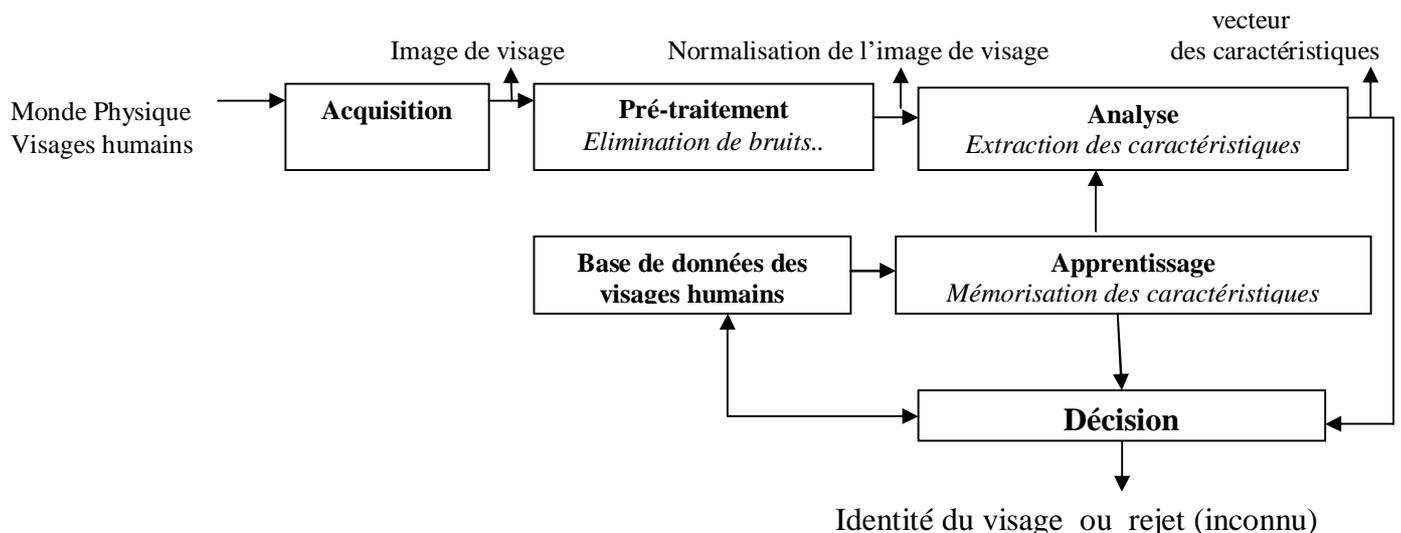


Figure 14 : Schéma général du processus de Reconnaissance des visages

- ✓ **Le monde physique**, il y a trois paramètres à considérer dans cet environnement :
 - L'éclairage, la variation de posture et l'échelle.
 - L'éclairage : L'image du visage est une matrice de niveaux de gris (C'est le cas dans la majorité des systèmes existants) représentant des intensités de lumière.
 - La variation de posture qui dépend de la position de la tête et de l'angle de prise de vue (vue de face, de profile, 45° ...).

∅ L'échelle : qui dépend de la taille de l'image et de la distance séparant l'individu de la caméra, une normalisation des distances est donc nécessaire.

La variation de l'un de ces trois paramètres peut conduire à une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents [22].

Une étude effectuée par Moses (1993) a montré que la variation de l'un de ces paramètres peut conduire à une distance (dans le sens mesure de similarité) entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents [23].

▼ **Le codage** : consiste en l'acquisition d'image et sa digitalisation, il comporte un risque de bruit et donne lieu à une représentation 2-D (la matrice des niveaux de gris) pour un objet 3-D (le visage). Peu de méthodes 3-D ont été proposées [24] , mais cela devrait changer dans le futur proche avec l'invention des Range-Cameras qui peuvent mesurer la topologie 3-D du visage.

Un problème émerge de façon automatique est celui de la définition et le choix de la représentation du visage , dans la littérature, il existe trois codages , les deux premiers décrivent l'apparence visuelle, le troisième utilise des vecteurs arbitraires ou aléatoire [25]

1- Codage numérique des visages : Les images des visages sont numérisées, et représentées par l'intensité lumineuse de chaque pixel .

2- Codage ou représentation géométrique : Les visages sont décrits par un ensemble de mesures des caractéristiques des visages telles que la longueur de nez, la distance entre les deux iris.....

▼ **Le prétraitement**, il faut éliminer le bruit par des techniques de traitement et de restauration d'images, et procéder à une localisation de la tête dans l'image. Cette opération est très complexe, et pour la contourner, beaucoup de chercheurs supposent que l'image contient un visage et que l'arrière plan est neutre. La restauration d'images ou l'élimination du bruit consiste à compenser les dégradations connues ou estimées et rétablir la qualité initiale de l'image.

Parmi les techniques déployées lors du prétraitement, on peut citer :

L'Echantillonnage :

La numérisation de l'image du visage est le résultat de l'échantillonnage de la fonction représentant cette image, il existe en général deux types d'échantillonnage [26]

1-Echantillonnage Spatial : définissant la finesse du maillage de discrétisation, c'est-à-dire le nombre de points par ligne qui donne la taille du point élémentaire.

2-Echantillonnage Dynamique : qui pour une dimension donnée du point élémentaire, définit la quantité d'énergie lumineuse en ce point.

▼ **L'Analyse** Dans cette étape (appelée aussi indexation, représentation, modélisations ou extraction de caractéristiques), il faut extraire de l'image les informations (les indices, ou points caractéristiques) représentant l'individu et qui vont être sauvegardés en mémoire secondaire (i.e. mémoriser l'individu en question) et qui seront plus tard utilisés dans la phase Décision. C'est au développeur de choisir quelles sont les informations utiles et comment les extraire, ceci revient à établir un modèle pour le visage. Ce modèle ou le choix des ces points caractéristiques doit respecter certains critères :

- **La discrimination** : les indices choisis doivent permettre de distinguer un individu parmi un groupe.

- **La non redondance** : comme dans les bases de données, il ne doit pas y avoir de redondance (inutile) d'information pour ne pas gaspiller l'espace de stockage et ralentir le système.

▼ **L'apprentissage** Il s'agit de mémoriser les représentations calculées dans la phase Analyse pour les individus connus. Cette sauvegarde dépend du modèle établi pour coder le visage, ce modèle va au fait imposer des structures de données qui peuvent aller d'un simple fichier séquentiel jusqu'à la création d'une base de données entière. Le temps d'accès à ces données constitue la plus grande partie du temps de réponse du système, il faut donc songer à utiliser des médias rapides.

▼ **La décision** : c'est la reconnaissance proprement dite, son rôle est d'identifier le visage test à partir de l'apprentissage réalisé.

La méthode de décision est souvent « exhibée » par l'apprentissage, ce qui veut dire que les critères utilisés pour la comparaison sont les même que ceux utilisés pour l'apprentissage.

Parmi les méthodes utilisées, certaines sont fondées sur la notion de proximité nécessitant le calcul d'une distance ou d'une probabilité de ressemblance avec les modèles

définis lors de la phase d'apprentissage, d'autres sont fondées sur l'analyse de la forme du visage, et essaient plutôt de vérifier une certaine cohérence entre les différentes sous-structures, les réponses fournies par cette étape peuvent être :

- Ø L'identité du visage (nom) en cas de bonne reconnaissance
- Ø Rejet du visage en cas d'incompatibilité de description avec les formes de références.
- Ø Visage inconnu en cas d'ambiguïté.

Dans la section 8 , nous présentons un aperçu sur les différentes méthodes déployées pour la reconnaissance des visages humains.

IV.5. Détection des visages

La détection des visages dans une scène ou image fixe est une tâche très simple pour le système visuel humain, aussi bien que pour un visage connu ou inconnu , On peut détecter un visage même dans les conditions les plus mauvaises : mauvais éclairage, longues distances, présence d'une partie seulement du visage ; s'emble-t-il que ce calcul est fait d'une manière massivement parallèle et probablement, l'être humain n'utilise pas toutes les informations relativement à la personne à reconnaître simultanément, le visage est perçu comme une totalité et non comme étant un ensemble de traits.

Pour un système automatique , la localisation et la détection d'un visage sont assez complexes, la résolution de l'image ainsi que les conditions dans lesquelles elle est prise sont très importantes pour la détection ; elle (la détection) est visualisée comme un outil très important pour un système de recherche qui consiste à donner une image simple et d'identifier toutes les régions d'images qui contiennent un visage indépendamment de sa position, orientation, et états tridimensionnels d'éclairage. [27]

La détection automatique de visages humains dans les séquences vidéo est généralement utilisée dans les systèmes temps réel. Elle nécessite non seulement la localisation de visage mais aussi le suivi du visage dans une séquence d'images.

Le but des systèmes de détection des visages est de pouvoir détecter des visages avec de bas taux d'erreurs et de donner les paramètres d'évaluation qui correspondent à la position, l'orientation, et la pose des visages avec une précision raisonnable , chacune de ces prévisions est associée à une mesure de certitude, on considère que les prévisions avec une mesure de certitude au-dessus d'un certain seuil sont rejetées, un visage est considéré être détecté quant au moins une prévision est dans une précision raisonnable de paramètres vrais.

Beaucoup de systèmes de détection automatique des visages ont été proposés se basant sur plusieurs principes et concepts tels que la couleur de la peau et les SVM dans le

système **This-is** proposé par [27], ainsi que le système **Where-is** proposé par [28], d'autres sur les propriétés chromatiques, le domaine de détection automatique des visages reste également très fertile, quoique des résultats très prometteurs ont été réalisés .

IV.6. Extraction des paramètres du visage

Avant de procéder à l'extraction des caractéristiques, une question à laquelle, beaucoup de chercheurs essayent de répondre : « combien de traits est-il nécessaire de prendre en considération pour différencier les visages et quels sont ces traits ? ». Pour cela nous disposons de quelques suggestions [29] disant que l'importance des traits dans un visage décroît de haut en bas, les yeux sont par conséquent d'une importance particulière dans la reconnaissance, car , parait-il que les iris sont les deux points uniques qui ne changent jamais de position quelque soit l'expression faciale, [30] avance que neuf points bien choisis sont suffisants pour identifier un visage dans une base de 80 personnes, mais que pour des bases plus large, le nombre doit être plus grand, typiquement de 35 à 45 points peuvent être localisés.

Donc cette étape est parmi les plus cruciales dans le processus de reconnaissance, vu que la décision (la classification) se basera sur la pertinence et l'exactitude avec lesquelles ces caractéristiques ont été extraites ,donc, le principe de l'extraction des primitives du visage est celui d'en effectuer la description c'est-à-dire d'extraire tous les objets constitutifs de la forme (visage) , de faire des mesures sur ces objets afin d'en retirer les paramètres pertinents permettant de les identifier et de donner leurs situations relatives.

Plusieurs méthodes d'extraction des caractéristiques du visage ont été élaborées dans la littérature citons parmi elles, celles qui se basent sur les contraintes suivantes :

- Ø La symétrie bilatérale du visage
- Ø Chaque visage possède :
 - § Deux yeux de part et d'autre.
 - § Un nez.
 - § Une bouche.
 - § Un menton.

avec la même disposition, ce qui rend la classification plus délicate, mais facilite du moins l'extraction.

Parmi les méthodes d'extraction des caractéristiques faciales citons :

IV.6.1 L'intégrale de projection verticale et horizontale [31]

Soit l'image du visage $I(x,y)$, L'intégrale de projection verticale dans le rectangle $[x_1,x_2]*[y_1,y_2]$ est défini par :

$$V(x) = \sum_{y=y_1}^{y_2} I(x, y) \quad \text{I.1}$$

L'intégrale de projection horizontale dans le rectangle $[x_1,x_2]*[y_1,y_2]$ est défini par :

$$H(y) = \sum_{x=x_1}^{x_2} I(x, y) \quad \text{I.2}$$

Ces projections aident généralement dans la localisation des traits , seulement il est préférable de limiter la fenêtre de travail afin d'éviter les erreurs , deux images de contours sont alors générées : horizontale (IH) et verticale (IV) :

- A partir de (IH) ,on tire les frontières du visage.
- A partir de (IV), on tire le sommet de la tête, les yeux, le bas du nez et la bouche.

Une fois les yeux localisés(généralement dans l'étape de normalisation), la localisation des autres traits devient plus simple.

▼ La Bouche et le Nez :

En premier lieu,on estime leur position par la recherche des pics dans la projection horizontale (IH) (nez), et des vallées dans la projection horizontale (la bouche), cette évaluation se fait selon la proéminence du pic ou de la vallée, ceux dont l'évaluation est la plus grande sont utilisés comme position verticale du nez et de la bouche.

Le nez est délimité horizontalement par la recherche des pics dans la projection verticale (IV) , la bouche est localisée par un seuillage de cette projection à la valeur moyenne.

De façon analogue, sont déterminés les sourcils,l'ensemble des traits sont alors normalisés et extraits .

IV.6.2 La méthode des moments invariants : [32]

Cette méthode introduite par Hu, Dudani et al est utilisée pour extraire les informations géométriques de l'image du visage en entrée représentée par une fonction $f(x,y)$, en calculant une séquence de nombres finis m_{pq} appelés « **moments** »

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad \text{I.3}$$

Pour une image binaire, ces moments seront définis par :

$$m_{pq} = \sum x^p y^q f(x, y) \quad \text{I.4}$$

Ces moments calculés sont utilisés comme des caractéristiques des visages à reconnaître, cette méthode comporte de nombreuses variantes, les plus connues sont celles basées sur les moments réguliers, les moments de Zernik.

IV.6.3 La transformée de symétrie généralisée :

C'est une méthode assez robuste et efficace proposée par Reisfeld [33], elle offre un outil puissant de détection des traits de symétrie [34], où l'image est déformée en une image standard en utilisant les points détectés, son principe est le suivant :

On commence par une edge map, où chaque pixel lui est assigné :

- Une magnitude M , qui estime la probabilité qu'il est une configuration spatiale symétrique des tranches qui l'entourent.
- Une orientation α , qui pointe dans la direction de l'axe principal de la symétrie qui est autour de ce pixel.

Par exemple, à un pixel qui se trouve au centre d'une surface circulaire, elliptique ou rectangulaire, entouré de tranches, doit être assignée une grande valeur de M , ceci résulte dans le map de symétrie, où chaque pixel a une valeur, et les plus hauts pics de symétrie doivent être détectés, donc, la détection de traits faciaux est réalisée en se basant sur la symétrie faciale, ce traitement diminue la dimensionnalité du problème.

D'autres méthodes sont utilisées pour la détection des traits faciaux, telles que celle employée pour générer des descripteurs de Fourier, des lignes du périmètre du squelette de la forme, et les modèles auto-régressifs [48].

D'autres travaux comme ceux de [61] [49], où ils présentent une autre approche d'extraction des primitives composée de deux phases, la première consiste à localiser les

yeux ; Le nez et la bouche seront durant la seconde recherchés en utilisant la géométrie du visage, son principe est le suivant :

▼ Localisation des yeux

L'existence d'une tâche de réflexion spéculaire due aux sources lumineuses frontales sur les iris, permet de localiser l'œil en utilisant l'intensité lumineuse de chaque pixel.

En faisant un balayage de l'image de haut en bas, ils ont pu faire ressortir les positions locaux de la fonction d'intensité de niveau de gris, ce qui permet de localiser les positions des points (pixels) possédant un maximum d'énergie, c'est-à-dire dont l'intensité est maximale. D'où, la détection des positions des iris, et tracer la droite qui les relie.

▼ Localisation du nez et de la bouche

Pour localiser le nez et la bouche, ils ont utilisé ensuite l'information spatio-temporelle de façon à caractériser les points ayant enregistré la plus forte variation d'intensité au cours de la séquence. L'image I_{som} ainsi obtenue vérifie :

$$I_{som} = \sum_{i=0}^{i=n} \sqrt{\left(\left(\frac{\partial I(x, y, t)}{\partial x} \right)^2 + \left(\frac{\partial I(x, y, t)}{\partial y} \right)^2 + m \left(\frac{\partial I(x, y, t)}{\partial t} \right)^2 \right)} \quad \text{I.5}$$

$I(x, y, t)$: L'intensité du niveau de gris.

x, y : Les composantes spatiales.

t : La composante temporelle.

m : Un terme de normalisation dépendant de l'échantillonnage temporel de la séquence et du mouvement du visage.

n : Le nombre des images acquises.

La localisation du nez et de la bouche consiste à chercher le long de la médiatrice passant par les iris, les deux régions ayant les plus fortes valeurs sur l'image de I_{som} .

Cette méthode permettait de localiser cinq points dans le visage : les centres des iris, le bout du nez, le point désignant le bout de la lèvre supérieur, le centre de la médiane passant par les iris constituait le cinquième point, ces points permettait de dégager huit distances qui constituait le vecteur des caractéristiques.

IV.7. Les méthodes de reconnaissance des visages

IV.7.1 Principe

L'identification des visages, qui est l'axe principal de notre étude et qui entre dans le cadre de la vision par ordinateur, est toujours un champ de recherche très fertile et dans lequel il y aura toujours de nouveaux résultats et de nouvelles théories capables de donner à la machine des capacités d'identification proches de celles de l'homme [49].

Dans le système visuel humain, l'identification est *le processus d'association du nom au visage*. Une identification correcte signifie que plusieurs images d'un même visage sont identifiées en tant qu'une seule.

Il est important de noter qu'il est difficile de décrire complètement un visage néanmoins, les traits faciaux et leurs relations géométriques jouent un rôle très important dans le processus d'identification. Cependant, un visage n'est pas une simple collection de traits, mais plutôt une synthèse de cette collection qui doit être établie pour une bonne identification.

Un autre problème que nous devons en tenir compte dans un système de reconnaissance automatique, est celui de pouvoir différencier entre le visage à identifier (dans l'image) et les visages connus dans la base des données. Pour cela, les visages connus et éventuellement leurs traits doivent être stockés dans une base de données

De plus, les êtres humains sont constamment en mouvement, même quand ils sont assis, ils remuent leurs corps pour s'ajuster, ils clignent des yeux, ils regardent au tour d'eux..., de ces constats, la majorité des systèmes supposent que l'image contient un visage et que l'arrière plan est neutre, Les autres systèmes qui n'ont pas cette limite [50] parviennent à détecter la position de la tête moyennant un certain overhead en temps de calcul.

Se focalisant sur l'aspect du visage comme un tout, ou une collection de traits, les méthodes de reconnaissance automatique de visage peuvent être répertoriées en deux catégories détaillées ci-dessous.

IV.7.2. Les Méthodes Existantes

Comme citer ci-dessus les systèmes de reconnaissance des visages humains peuvent être répertoriés en deux catégories, à savoir, les approches basées pixel appelées également méthodes globales, et les approches basées image dites également méthodes analytiques ou à traits, le diagramme I.3 illustre les ces deux familles et leurs sous- catégories.

IV.7.2.1. Les méthodes géométriques

On les appelle aussi les méthodes à traits, à caractéristiques locales, ou analytiques, ces méthodes se basent sur l'aspect image.

Principe : consiste à extraire la position relative des différents éléments du visage tels que le nez, la bouche, les yeux, les sourcils et la forme du contour facial afin de restreindre l'espace des données initiales, les visages sont alors modélisés et classés sur la base de distances et angles entre points caractéristiques. La configuration globale du visage peut être décrite par un vecteur de données numériques représentant la taille et la position des traits faciaux les plus importants, ensuite comparer ce vecteur avec ceux de la base de données en utilisant une méthode de comparaison appropriée.

Différentes méthodes statistiques, inspirées des algorithmes classiques de reconnaissance des formes, peuvent être alors utilisées pour discriminer les visages à partir de ces mesures [49] [22].

Ce modèle correspond à la manière avec laquelle l'être humain perçoit le visage, c'est à dire, à nos notions de traits de visages et de parties composites.

▼ **Etat de l'art des méthodes géométriques :**

[22] a développé un modèle géométrique du visage (**Who-is**) basé sur un ensemble de points caractéristiques au nombre de 12 extraites manuellement à partir de l'image du visage, la procédure de reconnaissance consistait à calculer les K plus proches voisins de l'individu test dans le sens de la distance city-Block ; Ils ont réalisé un taux de 86% de reconnaissance correcte sur un échantillon de 100 personnes.

En résumé, les approches géométriques dépendent étroitement de la performance des algorithmes de localisation et d'extraction des caractéristiques, cependant, ces dernières ne fournissent pas un degré élevé d'exactitude et exigent un temps de calcul considérable, une autre difficulté éprouvée quand il s'agit de prendre en considération plusieurs vues du visage, reste également du choix et de la pertinence psychologique des caractéristiques et leurs sélection.

▼ **Les avantages** de ces méthodes sont :

- Elles prennent en compte la particularité du visage en temps que forme naturelle à reconnaître, en exploitant les résultats des recherches en neuropsychologie et psychologie cognitive sur le système visuel humain.

- Elles sont simples à implémenter.

▼ **La difficulté** éprouvée quand il s'agit de :

- Prendre en considération plusieurs vues des visages ainsi que le manque de précision dans la phase "extraction" des points constituent.
- Temps de calcul considérable.

L'un des premiers travaux en reconnaissance automatique des visages utilisant les caractéristiques géométriques était réalisé par [51], l'auteur a proposé la description des profils faciaux tels que des courbes, trouver leur norme, ensuite classer les nouveaux profils par leurs dérivations par rapport à la norme, cette classification était multi-modal c'est-à-dire résultant d'un vecteur de mesures indépendantes qui peut être comparé avec d'autres vecteurs de la base, son système a réalisé une performance de 75% sur une base de 20 personnes utilisant deux images par personnes, une comme modèle, l'autre comme image test.

[52][53] ont montré qu'une approche géométrique utilisant une extraction manuelle des caractéristiques procure des résultats de loin meilleurs que ceux où l'extraction est automatique. [54] a automatiquement extrait un ensemble de caractéristiques géométriques, telles que la longueur et la largeur du nez, la position de la bouche, la forme du menton ; il en a extrait 35 constituant son vecteur de caractéristiques, la reconnaissance était par la suite effectuée par un classifieur Bayésien ; Son approche a atteint un taux de 95 % dans une base de données de 47 personnes.

Le système automatique de Vezjak [55] traite les vues de face pour des visages de personnes masculines avec un arrière plan neutre et à une distance connue de la caméra. Il effectue une approximation des parties du visage par des formes géométriques avec des techniques de détection de contour (opérateur de Hough) pour localiser la tête et extraire plus de quarante points utilisés dans le calcul d'un vecteur caractéristique. Ce vecteur est composé de dix-neuf paramètres qui sont des distances exprimées en pixels et des angles en degrés. L'identification consiste à chercher le vecteur enregistré dans la base dont la distance euclidienne qui le sépare du vecteur test est minimale.

Le système effectue une identification correcte, mais le test était limité à 24 photos pour 12 personnes (2 photos par personne, l'une pour l'apprentissage l'autre pour le test).

IV.7.2.2. Les Méthodes globales

Les méthodes globales se basent principalement sur l'information pixel de l'image du visage, ces dernières nécessitent peu de connaissances préalables sur la structure du visage, contrairement à la première donc, cette deuxième classe regroupe les méthodes qui mettent en valeur les propriétés globales de la forme [30]. Ici, le visage est traité comme un tout (dans sa totalité) plutôt que d'avoir affaire à des parties telles que les yeux, la bouche, etc.

Certains travaux s'inspirent des méthodes de traitement d'image liées à la mise en correspondance de régions; les mesures utilisées pour identifier les visages sont alors obtenues par des calculs de corrélation entre les régions; d'autres méthodes utilisent une analyse en composantes principales réalisée à partir de l'information pixel, pour rechercher une base de représentation permettant de discriminer les visages [57]

Les techniques globales travaillent le plus souvent pour la classification des prises frontales du visage (vue de face); cependant, elles ne sont pas robustes contre la variation de pose, ceci parce que les caractéristiques globales sont extrêmement sensibles à la rotation du visage; pour éviter ce problème, des chercheurs comme [60] ont introduit une étape d'alignement avant la classification du visage.

Aligner une image de visage test avec une image de visage de référence exige un calcul de correspondance entre ces deux dernières, cette correspondance est déterminée par un nombre restreints de points éminents dans le visage tels que : le centre des yeux, les narines, les coins de la bouche; se basant sur ces correspondances, l'image de visage d'entrée est fléchiée à l'image de visage de référence.

IV.7.2.2.1 Les visages propres (ACP) :

C'est l'approche la plus investit en reconnaissance de visages, connue également comme étant l'extention de Karhunen-Loève [56], vecteurs propres, ou bien l'Analyse en Composantes Principales ACP.

Son principe est le suivant : trouver les vecteurs et les valeurs propres de la matrice de covariance C à partir de l'ensemble des images d'apprentissage $\{T_1, T_2, \dots, T_M\}$,

$$C = \frac{1}{M} \sum_{i=1}^M (T_i - \Psi)(T_i - \Psi)^T$$

Où ψ est le visage moyen; Par projection des images dans l'espace des visages formé par les vecteurs propres, on peut calculer exactement leurs distance euclidienne, l'image de

l'ensemble de l'apprentissage ayant la distance minimale qui la sépare de l'image test sera prise en compte pour la reconnaissance .

Les étapes suivantes résument le processus de reconnaissance [57]

1. Initialisation : Acquérir les M images de visages constituant l'ensemble d'apprentissage et calculer les visages propres qui définissent l'espace des visages.
2. Pour l'image d'un nouveau visage, calculer un ensemble de poids basé sur l'image en entrée et les M visages propres par projection de cette image sur chaque visage propre.
3. Déterminer si l'image contient un visage en vérifiant si elle est suffisamment proche de l'espace des visages.
4. Si l'image contient un visage, on le classe selon les poids précédemment calculés comme étant une personne connue ou inconnue.

Les références [58] [56] ont utilisé l'ACP pour représenter efficacement les images des visages ; ils ont déduit que n'importe quelle image de visage peut être approximativement reconstruite à partir d'une collection de poids pour chaque visage et image standard de visage, les poids décrivant chaque visage sont obtenus par projection de l'image du visage sur l'image propre.

Mathématiquement parlant, les visages propres sont les composantes principales de la distribution des visages ; les vecteurs propres sont ordonnés pour représenter différentes grandeurs de la variation parmi les visages.

Chaque visage peut être représenté donc par une combinaison linéaire des visages propres, il peut également être rapproché en utilisant uniquement le meilleur vecteur propre ayant les plus grandes valeurs propres ; les M meilleurs visages propres construisent un espace de dimension M appelé l'espace des visages.

IV.7.2.2 Les réseaux de neurones

L'attraction vers l'utilisation des réseaux de neurones peut être due à l'aspect de la non linéarité dans ces derniers, l'étape d'extraction des caractéristique serait plus efficace que la méthode linéaire de Karhunen-Loève.

La première technique utilisant les réseaux de neurones artificiels (ANN) pour la reconnaissance de visages humains était un réseau à une seule couche adaptative nommé WISCAD [59] de Stonham (1980) traitait les images binaires des visages de 16 individus (vue de face et avec variation d'expressions) et qui contenait un sous-réseau séparé pour chaque

individu ; WISCAD était capable de reconnaître avec un certain succès l'identité et l'expression du visage, mais lors de l'apprentissage, près de 400 images étaient utilisées pour les 16 individus.

[35] a proposé un réseau de neurones hybride qui combinait des échantillons locaux de l'image, une carte auto-organisatrice et un réseau de neurone convolutionnel ; La carte auto-organisatrice fournit une quantification des échantillons de l'image dans un espace topologique où les entrées qui sont tout près les unes des autres dans l'espace originel, le seront dans l'espace des sorties, de ce fait, la carte auto-organisatrice fournit une réduction des dimensions et une invariance pour des changements mineurs dans les échantillons de l'image du visage ; appliquée sur la base de données ORL , le temps de classification était moins d'une seconde, mais l'apprentissage dépassait les 04 heures , le taux de reconnaissance était de 96,2 %.

[36] a utilisé un réseau de neurones basé sur une décision probabiliste qui hérite de la structure modulaire de son prédécesseur [37](un réseau de neurones basé sur la décision) ; ce réseau a prouvé son efficacité pour :

- 1- La détection du visage dans l'image.
- 2- Localisation des yeux et génération du vecteur des caractéristiques.
- 3- La reconnaissance faciale des visages .

Ce type de réseau n'a pas une topologie entièrement connectée, au lieu de cela , il divise le réseau en k sous-réseaux ; où chaque sous-réseau est dédié pour la reconnaissance d'une seule personne dans la base de données , ils utilisent la fonction du Gaussien pour l'activation de ces neurones ;

Le schéma d'apprentissage de ces réseaux se produit en deux phases :

- 1- Chaque sous-réseau est entraîné par son propre visage .
- 2- Dans cette phase appelée :Décision basée sur l'apprentissage , les paramètres du sous-réseau peuvent être réentraînés par des échantillons particuliers provenant des autres classes , ce schéma n'utilise pas tout les échantillons pour le réapprentissage, il emploie uniquement les modèles mal classés ; [38] a rapporté que ce système a la capacité de reconnaître correctement plus de 200 personnes , cependant , lorsque le nombre de personnes augmente , le système perd ces capacités et le coût de calcul devient assez important.

Les systèmes biométriques utilisant les réseaux de neurones basés sur décision probabiliste ont le mérite des réseaux de neurones et des approches statistiques , aussi le principe de calcul distribué est relativement simple et facile à implémenter .

En résumé, les approches basées sur les réseaux de neurones sont assez perspicaces mais rencontrent des problèmes lorsque le nombre de classes devient important c'est à dire que l'extension à des problèmes plus large n'est pas clair , et l'apprentissage est assez long.

IV.7.2.2.3 Les templates matching (Appariement de modèles)

Une version simplifiée des templates matching est que l'image test est représentée par un vecteur bidimensionnel des valeurs d'intensités ; celui-ci est comparé en utilisant une métrique appropriée telle que la distance euclidienne avec un seul template représentant le visage entier, un template est un masque qui se compose par exemple des yeux, de la naissance des cheveux, des sourcils , du nez et de la bouche ; Il existe d'autres versions des templates matching plus sophistiqués pour la reconnaissance de visage ou plus d'un template du visage peuvent être utilisés pour représenter le visage.

Dans [[31](#)] , Bruneli et Paggio choisissent automatiquement un ensemble de quatre templates de caractéristiques , à savoir, les yeux, le nez, la bouche et le visage entier , ceci pour tout les visages disponibles, ils ont comparé leur technique avec une approche géométrique sur une même base de données qui contenait 188 images de 47 individus ; Les templates matching étaient de loin supérieur , un taux de reconnaissance de 100% a été réalisé contre 90% pour l'approche géométrique , de plus les templates étaient plus facile à implémenter ; Ils ont déduit que les approches basées sur les templates matching seraient plus logiques que celles basées sur le templates des caractéristiques, néanmoins , un inconvénient éprouvé à leurs encontre est leur complexité de calcul , et le problème de base qui se pose est le nombre de templates requis.

IV.7.2.2.4 Les Machines à vecteurs de support SVM

Les SVM sont des algorithmes de classification récemment développés par Vapnick [[39](#)]. L'idée de base des SVM est d'utiliser des fonctions noyaux et de projeter les données de l'espace d'entrée (non linéairement séparables) dans un espace de plus grande dimension appelé espace de caractéristiques , de façon à ce que les données deviennent linéairement séparables .Ces algorithmes tentent de minimiser les erreurs d'apprentissage tout en maximisant la marge séparant les données de classes. Le comportement du SVM est conditionné par le type de noyau utilisé dans la projection et les valeurs accordées à ces paramètres.

Les SVM peuvent également être vu comme une manière d'apprentissage pour les réseaux de neurones polynomiaux ou ceux de fonctions à base radiale.

L'application des SVM dans le problème de reconnaissance de visage a été proposé récemment dans [40] où après extraction manuelle des caractéristiques, des fonctions de discrimination entre chaque paire de vecteurs des caractéristiques sont apprises par les SVM, ensuite l'ensemble de test de disjonction est entré au système pour la reconnaissance, ils ont appliqué leur approche sur une la base de données ORL, l'erreur minimale était de 8,78% qui était nettement inférieur à 14,15% fournit par l'application de l'approche ACP sur la même base.

En [41], le problème de reconnaissance de visage a été formulé comme étant un problème dans un espace de différence qui modélisaient les dissimilarités entre deux images faciales. Dans plusieurs espaces différents, ils ont formulé le problème de reconnaissance de visage comme un problème à deux classes :

- 1- Les dissimilarités entre les visages d'une même personne.
- 2- Les dissimilarités entre les visages de personnes différents.

Par modification de l'interprétation de la décision, la surface générera une métrique de similarité entre les visages. L'algorithme basé sur les SVM a été comparé à l'approche des visages propres ACP sur un ensemble assez difficile d'images de visage de la base de données FERET, la performance d'identification était de 77.78% pour les SVM contre 54% pour l'ACP ; le taux d'erreur égalé les 7% pour les SVM contre 13% pour l'ACP.

[42] a présenté une technique basée sur les composantes SVM et deux autres techniques globales pour la reconnaissance de visage, il a évalué leur performance en ce qui concerne leur robustesse contre les variations de pose ; Le système basé sur les composantes détecte et extrait un ensemble de dix composants faciaux et les disposaient dans un vecteur de caractéristiques qui sera classifié par un SVM linéaire ; Dans les deux systèmes globaux, le visage entier est détecté et extrait de l'image, il est utilisé comme entrée pour les classifieurs :

- Ø Le premier système globale est composé d'un SVM unique pour chaque personne de la base de données.
- Ø Dans le second, une base de données pour chaque personne est regroupée et entraînée sur un ensemble de classifieurs SVM.

Les systèmes ont été testés sur une base de données de 8593 images de visage à un seul niveau de gris incluant des visages avec une rotation de 40° ; dans tous les cas de figures, les systèmes basés sur les composantes SVM ont surpassé les systèmes globaux .

D'autres part, [43] et son équipe ont étudié les SVM dans le contexte de la vérification, dans leurs études , ils ont conclu que l'approche SVM est capable d'extraire de l'information discriminante à partir de données d'apprentissage , et ceci est la raison principale de sa performances .

En conclusion , l'approche SVM est capable d'extraire de l'information discriminantes et pertinentes à partir des données d'une manière entièrement automatique ; elle peut également faire face aux variations d'illumination , le rôle majeure dans cette caractéristique revient à la capacité des SVM d'apprendre les décisions des frontières non linéaires.

IV.7.2.2.5 Les modèles de markov cachés HMM

La modélisation stockastique du vecteur non stationnaire basée sur les modèles de Markov cachés ont réalisé un exploit dans les applications de la reconnaissance de la parole .

[44] a appliqué cette méthode pour la reconnaissance de visage humain, où les visages sont intuitivement divisés en régions telles que : les yeux, le nez, la bouche...., qui peuvent être associées à des états du Modèle de Markov Caché ; et vu que les HMMs requièrent une séquence d'observations uni-dimensionnelles, et que les images de visages sont bidimensionnelles, l'auteur a proposé la conversion de ces dernières soit en une séquence temporelle ou spatiale unidimensionnelle .

En [45], une séquence d'observations spatiales est extraite de l'image en utilisant une technique d'échantillonnage ; chaque image de visage est représentée par la suite par un vecteur uni-dimensionnel de pixels des séries d'observations ; chaque vecteur est un bloc de L lignes , et il y 'aurait M lignes chevauchées entre la suite d'observations, premièrement , une image test inconnue est échantillonnée de la séquence d'observations, puis elle est mise en correspondance avec les HMM dans le modèle de la base de données des visages (chaque HMM représente différents sujets) ; la correspondance ayant la plus grande probabilité est considérée comme la meilleure correspondance , et le modèle qui s'y rapporte indique l'identité du visage test , cette méthode a été appliquée sur la base de données ORL , elle réalisa un taux de reconnaissance de 87 %.

Deux inconvénients sont reprochés aux HMM ,

- 1 – Le temps d'apprentissage et celui de la classification est trop élevé.
- 2- Le choix des paramètres est basé sur l'intuition subjective.

IV.7.2.2.6 Les systèmes multi-classifieurs

Il existe principalement deux structure de combinaison :

1 – *Structure statique* : les données d'entrées ne sont pas impliquées dans la combinaison , elles sont des entrées pour tout les classifieurs, seules les sorties de chacun d'eux seront combinées suivant une règle appropriée (vote, sommation , produit)

2- *Structure Dynamique* : Les entrées sont directement impliquée dans le mécanisme de combinaison qui utilise une unité d'intégration pour ajuster le poids de chaque classifieurs.

[46] a présenté un système de reconnaissance de visage utilisant les capacités de généralisation des LVQ (Learning Vector Quantification), et un réseau de neurones de fonctions à base radiale pour la construction d'un modèle représentatif du visage à partir d'une variété de formes d'apprentissage avec différentes poses et expressions faciales ; l'auteur avance que la performance de son système est supérieur à celle de chaque classifieur pris seul avec les mêmes variations

[18] ont développé un comité de machines pour la reconnaissance de visage , qui regroupé cinq algorithmes à savoir les visages propres (ACP) , Les visages de Fisher , Elastic Graph Matching (EGM), Les machines à vecteurs de support (SVM) et les réseaux de neurones , ceci pour obtenir une décision unifiée avec une exactitude améliorée ; ils discutent un taux de reconnaissance de 98,8% sur la base de données des visages ORL , ce taux dépasse ceux réalisés par chacun des classifieurs pris individuellement et testé sur la même base .

Dans [47] , un système de classifieurs combinés qui était construit d'un ensemble de réseaux de neurones , ce dernier était basé sur la variation des paramètres de ces réseaux ; un algorithme d'amplification était utilisé pour produire une perturbation dans l'ensemble d'apprentissage en utilisant un MLP comme base du classifieur , le résultat final est combiné en utilisant une simple règle de vote majoritaire ; Ce système a réalisé un taux de reconnaissance de 99,5% sur la base de données YALE , et de 99,8 % sur celle d'ORL

IV.7.3 Avantages et limites des méthodes de reconnaissance des visages :

Les points forts et les limite de ces méthodes peuvent être résumés dans le tableau ci-dessous :

APPROCHES	AVANTAGES	LIMITES
Approches Géométriques	<ul style="list-style-type: none"> - Requier moins d'espace mémoire.. - Plus rapide et simple à implémenter 	<ul style="list-style-type: none"> -Dépendent étroitement de l'exactitude d'extraction des caractéristiques -Prise en charge de la variation de la pose,de l'éclairage..
Les visages propres	<ul style="list-style-type: none"> -Une représentation discriminante du visage . - Réduction de la taille du problème 	<ul style="list-style-type: none"> -Coût de calcul élevé . -Détermination du nombre de visages propres(base de projections)
Les templates Matching	Robuste pour localiser les régions	<ul style="list-style-type: none"> -Complexité de calcul -Le nombre de templates -Espace mémoire important
Les Réseaux de Neurones	<ul style="list-style-type: none"> - Temps de classification - Résistent aux bruits et au manque de fiabilité des données - Prise en compte des variations de poses, d'eclairage 	<ul style="list-style-type: none"> -Sur-apprentissage -Extension à des problèmes plus larges -Apprentissage long
Les HMMs	- Algorithme d'apprentissage très robuste	<ul style="list-style-type: none"> -Temps d'apprentissage et classification long. -Choix des états basé sur l'intuition subjective
Les Machines à Vecteurs de Support	-Capacité d'extraction des informations discriminantes et pertinentes du visage d'une manière entièrement automatique.	- Temps d'apprentissage long et algorithme d'apprentissage compliqué.

Tableau 2 : Avantages et Limites des méthodes de classification des visages.

IV.8. Principaux domaines d'utilisations

L'analyse de la forme du visage a souvent été utilisée au cinéma dans le monde de la fiction. Aujourd'hui cette technique commence à être utilisée de plus en plus dans le monde réel pour identifier les personnes dans les lieux publics. Citons quelques exemples :

- Ø Dans les casinos aux Etats-Unis, cette technique est utilisée pour l'identification des joueurs interdits ou les gangsters fichés.

- Ø Dans les aéroports, elle est utilisée pour l'identification d'individus indésirables sur le territoire.
- Ø Dans les stades où elle est utilisées pour le refoulement de voyous connus et dangereux pour la tranquillité des autres spectateurs.
- Ø Dans les centres commerciaux, pour repérer l'entrée de petits voleurs connus des responsables de la sécurité.
- Ø Le contrôle d'accès aux institutions défensives et gouvernementales
- Ø L'accès aux bases de données sensibles (secrets d'état....) , aux réseaux informatiques hautement confidentiels

9. Avantage de la biométrie

Biométrie : remplace des objets et des codes par des caractéristiques physiques de l'individu. C'est un moyen pour :

- ü Eviter le risque de perte ou de vol.
- ü La simplification d'usage (pas de mémorisation nécessaire).
- ü Eviter l'imitation car elle a des caractéristiques difficile.
- ü Une délégation des droits.
- ü Une signature biométrique puisqu'elle permet une identité du responsable de l'opération.

10. Conclusion

Une chose reste évidente est que le visage est une entité particulièrement exceptionnelle, par la richesse d'informations qu'elle véhicule, des informations allant des plus concrètes aux plus abstraites, donc, il reste un élément à explorer et à découvrir.

On peut conclure que malgré la pléthore de méthodes et d'approches élaborées et testées pour parvenir à la construction de système d'identification basé sur le visage , et les efforts déployés par les chercheurs dans plusieurs domaines afin de décoder ce mécanisme mystérieux de la perception et du raisonnement humain , le domaine de la reconnaissance automatique des visages humains reste une terre toujours fertile et à explorer .

Chapitre IV

Conception proposée

1. Introduction

Nous avons présenté dans les chapitres précédents un état de l'art sur les domaines qui touchent à notre travail. C'est à partir de ce chapitre que commence notre contribution. Rappelons que notre objectif est la conception et l'implémentation d'un système basé fusion des données Pour Un Raisonneur Par Région de visages humains avec la comparaison entre les différents approches proposés (k-ppv avec tout le visage, k-ppv pour chaque région du visage, k-ppv pour les yeux avec le nez, k-ppv pour les yeux avec la bouche, k-ppv pour le nez avec la bouche).

2. Motivation

2.1. Pourquoi la reconnaissance des visages

La reconnaissance des visages est l'une des techniques de biométrie la plus utilisée, ceci est dû à ses caractéristiques avantageuses dont on peut citer :

- ü Disponibilité des équipements d'acquisition et leur simplicité : les systèmes de sécurité et de contrôle, basés sur la reconnaissance des visages, peuvent utiliser une simple caméra ou appareil photo numérique en plus d'un simple ordinateur personnel.
- ü Passivité du système : un système de reconnaissance des visages ne nécessite aucune coopération de l'individu, du genre : mettre le doigt ou la main sur un dispositif spécifique ou parler dans un microphone. En effet, la personne n'a qu'à rester ou marcher devant une caméra pour qu'elle puisse être identifiée par le système.
- ü En plus, cette technique est très efficace pour les situations non standard, c'est les cas où on ne peut avoir la coopération de l'individu à identifier, par exemple lors d'une arrestation des criminels.

Certes que la reconnaissance des visages n'est pas la plus fiable comparée aux autres techniques de biométrie, mais elle peut être ainsi si on utilise des approches plus efficaces en plus du bon choix des caractéristiques d'identification représentant le visage en question.

2.2. Pourquoi les k plus proches voisins

L'approche par les k plus proches voisins (k-PPV) possède au moins deux avantages : elle est très simple et elle ne nécessite pas d'apprentissage. La modélisation utilisée pour la classification est simplement la base d'apprentissage elle-même qui constitue l'ensemble des connaissances du système. Il n'y a donc pas d'architecture à proprement parler. La règle de décision utilisée consiste à attribuer à une entrée e la classe représentée majoritairement dans son voisinage. Ce voisinage est défini par les k plus proches individus de e parmi ceux se trouvant dans la base d'apprentissage. Les seuls paramètres à déterminer sont donc k ainsi que la métrique utilisée pour comparer les individus et trouver les plus proches voisins. Dans plusieurs cadres applicatifs, cette approche a donné des résultats intéressants. Cependant elle possède plusieurs inconvénients. Son efficacité dépend directement de la pertinence de la base d'apprentissage et notamment de sa densité dans les différentes régions de l'espace des données. La présence de bruit (données mal étiquetées) est aussi fortement handicapante. Autre inconvénient : la base d'apprentissage entière doit être stockée, ce qui nécessite en général une place mémoire importante. Enfin la recherche des plus proches voisins est coûteuse et cela d'autant plus que la métrique utilisée est complexe.

3. présentation de Notre Système

Les enquêteurs sont les utilisateurs typiques de ce système, il peut les assister dans l'identification de suspects en disposant d'une base d'images des visages des personnes recherchées. Le système reçoit en entrée l'image du visage de 3 façons possibles :

- Par appareil photo numérique (présence de l'individu à reconnaître).
- Par scanner (photographie de l'individu).
- Ou à partir d'un fichier graphique.

L'image acquise est ensuite affichée sur l'écran pour que l'opérateur puisse à l'aide de la souris localiser les points caractéristiques choisis dans le modèle. Ensuite, Le système effectuera la reconnaissance et affichera le résultat obtenu.

Notre système possède les caractéristiques suivantes :

- Ne traite que la vue de face du visage. (Ne gère pas le problème de variation de posture).
- Invariant à la présence de moustaches et de barbes, et indifférent aux expressions faciales (tant que la bouche n'est pas totalement couverte ni trop ouverte).
- Invariant au port de chapeau, casquette, foulard, de lunettes (tant que sa ne cache pas les yeux). Le maquillage n'a pas d'influence non plus.
- Pas de contraintes imposées à l'environnement (Changement d'éclairage est sans influence, et l'arrière plan peut être quelconque).
- L'image peut être en niveaux de gris ou contenir un nombre quelconque de couleurs.

4. La base des images utilisée :

Afin de monter la performance de notre application il convient notamment d'effectuer des tests sur une grande base d'image prise comme échantillon. A cet effet, nous avons choisi de travailler sur la base (Faces94) cette base d'images contient plus de 150 individus de sorte que chaque individu possède 20 images chacune d'elles représente une expression faciale différente. Parmi les vingt images de chaque individu, nous avons consacré la première image pour la phase d'apprentissage alors que les 19 restantes sont réservées pour la phase du test.

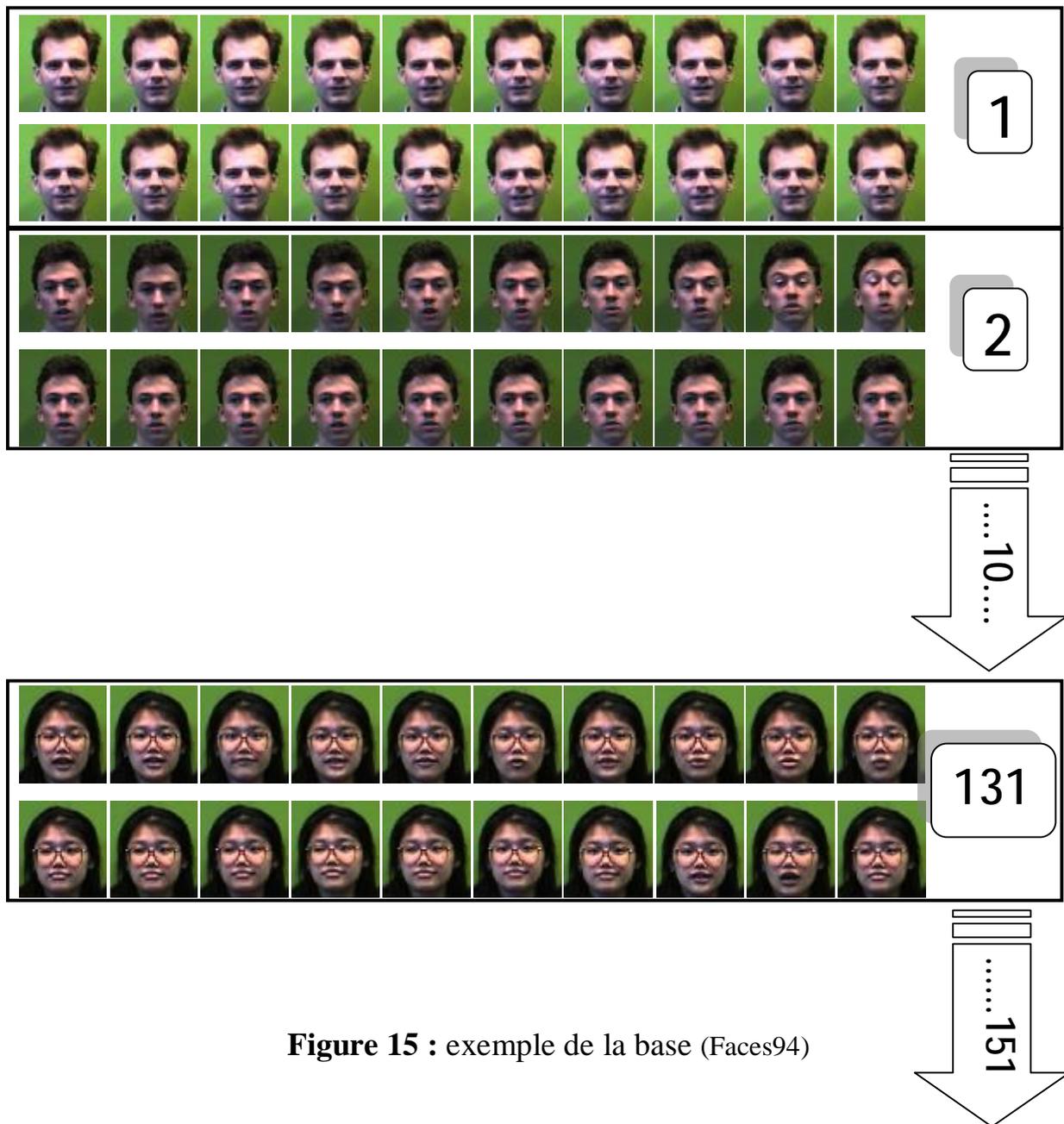


Figure 15 : exemple de la base (Faces94)

5. Système de reconnaissance de visages:

Nous présentons dans cette partie, la méthode utilisée pour effectuer la reconnaissance de visages. Rappelons pour commencer l'objectif à atteindre. On possède une image contenant un visage dont on veut identifier. Les points caractéristiques (à extraire) sont définis par un modèle géométrique. Où on a 12 points bien choisis qui sont suffisants pour identifier un visage, mais pour des bases de données plus larges le nombre de points caractéristiques doit être plus grand. Le choix des points caractéristiques est basé sur le fait que la région des yeux, celle du nez et celle de la bouche soient importantes pour l'identification, les points

qu'on utilise couvrent l'ensemble des trois régions, ils sont au nombre de 12 : Nous avons réutilisé 12 points [22], tel qu'illustré par la Figure 16 :

- Les quatre coins des yeux : A, B, C, et D.
- Le point S centre de gravité de A, B, C et D (milieu de la tête).
- Le point G à la base du nez.
- Les points extrême droit et extrême gauche du nez : E et F.
- Le point H milieu du contour supérieur de la lèvre supérieure.
- Les deux coins de la bouche I et J.
- Le M à l'intersection des droites (SG) et (IJ) représentant le milieu de la bouche.

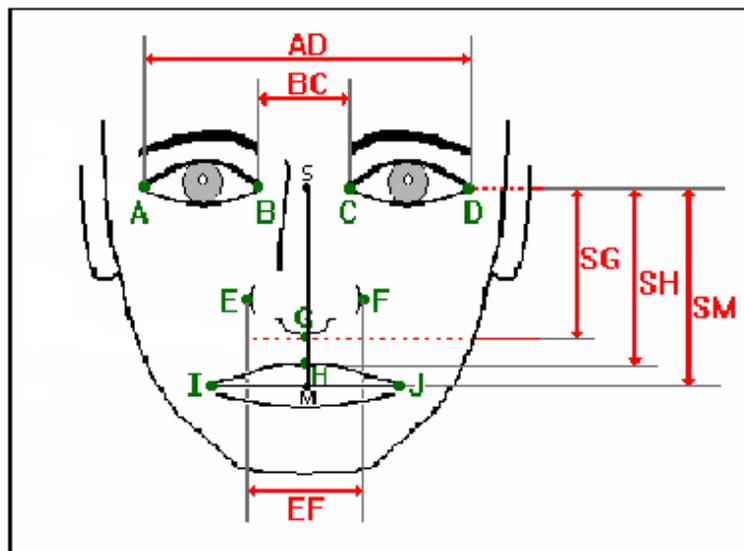


Figure 16. Modèle géométrique du visage

Les coordonnées des 12 points seront utilisées dans le calcul d'un vecteur caractéristique qui représente l'individu, les composantes du vecteur caractéristiques sont des distances normalisées entre les 12 points extraits. Les informations choisies pour modéliser le visage doivent être discriminantes, compactes et non redondantes, nous avons sélectionné 6 parmi les 66 distances possibles entre les 12 points, à raison de 2 distances pour chaque région du visage :

- Région des yeux : Généralement l'œil gauche a la même taille que l'œil droit et donc les points A et B sont symétriques aux points C et D par rapport à S, pour éviter la redondance d'information ($AB = CD$ et $AC = BD$) nous n'avons utilisé que les 2 distances AD et BC pour coder la région des yeux.

-
- Région du nez : La hauteur du nez est représentée par la distance SG, et la largeur par la distance EF.
 - Région de la bouche : Etant donné que la lèvre supérieure est liée à la mâchoire supérieure qui est immobile (contrairement à la mâchoire inférieure), on peut dire que la distance SH (voir Figure 2) est invariante, et de même pour la distance verticale SM puisque dans la plus part des cas la bouche n'est pas trop ouverte. Ainsi, dans sa totalité, le vecteur caractéristique est : $V[AD BC EF SG SH SM]$

Avec ce système on attribue une identité à un individu ou on le déclare inconnu. La méthode de reconnaissance adoptée est celle des k plus proches voisins (KPPV). La distance utilisée est la distance City-Block Avec la méthode des K-PPV il faut calculer la distance entre le vecteur caractéristique de la personne à identifier et chaque vecteur W_i d'un ensemble E qui contient les vecteur W_i de la base.

6. Architecture de Système

La Figure 17 montre le schéma bloc du système avec ses deux modes de fonctionnement : Apprentissage et Identification.

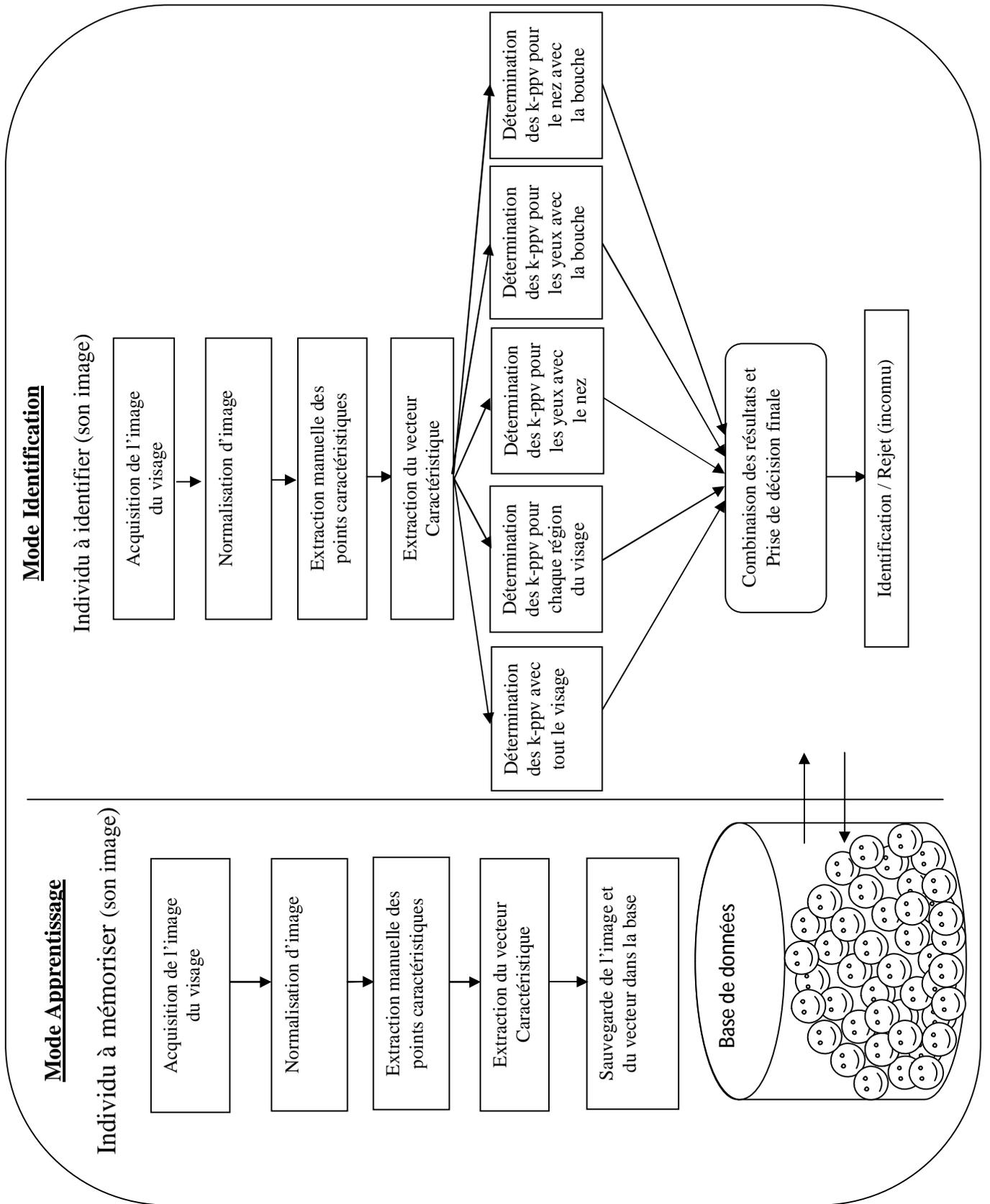


Figure 17. Schéma bloc du système

7. Un système d'identification

Les coordonnées des 12 points seront utilisées dans le calcul du vecteur caractéristique qui représente l'individu, et qui sera sauvegardé dans la base du système pour être utilisé plus tard dans l'identification.

Les composantes du vecteur caractéristiques sont des *distances normalisées* entre les 12 points extraits. Comme les informations choisies pour modéliser le visage doivent être discriminantes, compactes et non redondantes, nous avons sélectionné 6 parmi les 66 distances possibles entre les 12 points, à raison de 2 distances pour chaque région du visage :

- *Région des yeux* : Généralement l'œil gauche a la même taille que l'œil droit et donc les points A et B sont symétriques aux points C et D par rapport à S, pour éviter la redondance d'information ($AB = CD$ et $AC = BD$) nous n'avons utilisé que les 2 distances AD et BC pour coder la région des yeux.
- *Région du nez* : La hauteur du nez est représentée par la distance SG, et la largeur par la distance EF.
- *Région de la bouche* : Etant donné que la lèvre supérieure est liée à la mâchoire supérieure qui est immobile (contrairement à la mâchoire inférieure), on peut dire que la distance SH (voir Figure) est invariante, et de même pour la distance verticale SM puisque dans la plus part des cas la bouche n'est pas trop ouverte.

Ainsi, dans sa totalité, le vecteur caractéristique est :

$V [AD BC EF SG SH SM]. \quad [22]$

8. Méthode de reconnaissance

Avec notre système on attribue une identité à un individu ou on le déclare inconnu. La méthode de reconnaissance adoptée est celle des k plus proches voisins (KPPV). La distance utilisée est celle de la somme des valeurs absolues appelée aussi distance de Minkowski de degré 1 ou distance City-Block, elle est équivalente, aux autres distances (e.g. la distance euclidienne,...) mais elle nécessite moins de calculs (à ne pas confondre avec les distances AD, BC,...,SM).

Avec l'approche de fusion des données proposée, plusieurs méthodes de calcul des k-ppv sont mises au point pour optimiser la reconnaissance, c'est-à-dire une liste de k-ppv pour chaque méthode:

8.1. Détermination des k-ppv d'après tout le visage : afin de déterminer les k-ppv selon le (visage global) il faut calculer les 150 distances D_i (distances City-Block) entre le vecteur caractéristique de l'individu testé et chaque élément de l'ensemble de la base d'images, c'est-à-dire l'obligation d'intégrer toutes les 6 distances pour calculer la distance entre deux individus, contrairement aux autres méthodes proposés qui utilisent seulement une partie de vecteur caractéristique.

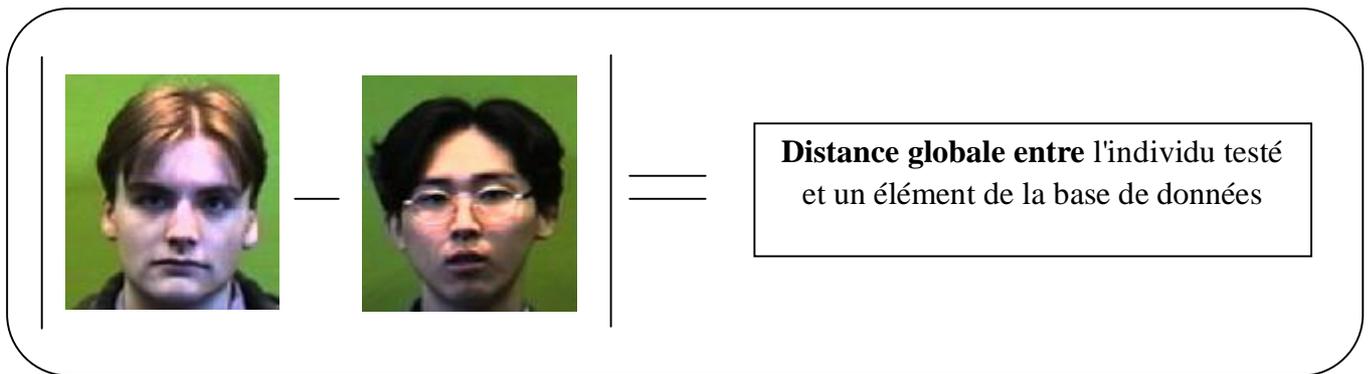


Figure 18 : Distance globale

le calcul des distances D_i (distances City-Block) entre le vecteur caractéristique $V(v_1, v_2, \dots, v_6)$ de l'individu testé et chaque élément $W_i(w_{i1}, w_{i2}, \dots, w_{i6})$ de l'ensemble E des individus de la BDD à l'identification :

$\sum_i |V_j - W_j^i|$ où W_j^i et V_j , $j=1..6$, sont les distances normalisées du modèle.

8.2. Détermination des k-ppv pour chaque région du visage :

8.2.1. Détermination des k-ppv pour les yeux : pour Déterminer les k-ppv pour les yeux il faut utiliser seulement les distances AD BC.

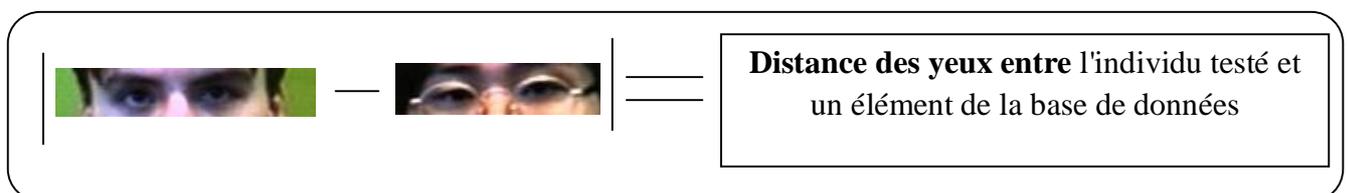


Figure 19 : Distance des yeux

8.2.2. Détermination des k-ppv pour le nez : pour Déterminer les k-ppv pour le nez il faut utiliser que les distances EF et SG

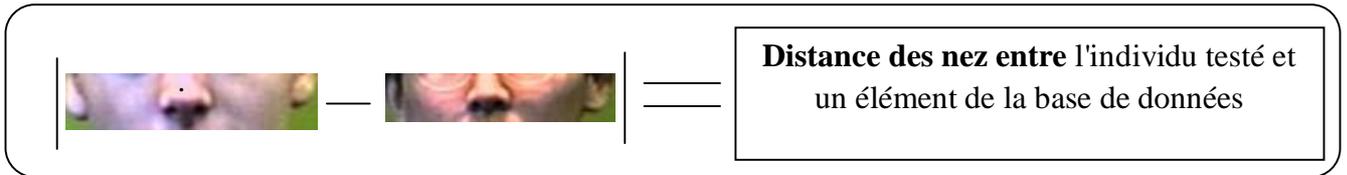


Figure 20 : Distance des nez

8.2.3. Détermination des k-ppv pour la bouche : pour Déterminer les k-ppv pour la bouche il faut utiliser que les distances SH et SM.

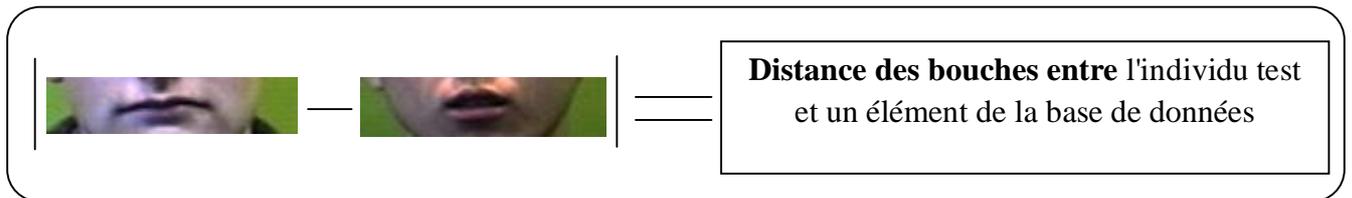


Figure 21 : Distance des bouches

8.3. Détermination des k-ppv pour les yeux avec le nez : pour Déterminer les k-ppv pour les yeux avec le nez on a utilisé que les distances AD, BC, EF, SG.

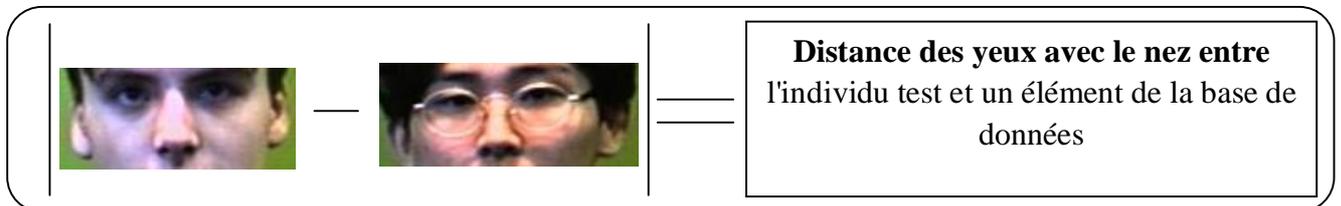


Figure 22 : Distance (yeux,nez)

8.4. Détermination des k-ppv pour les yeux avec la bouche : pour Déterminer les k-ppv pour les yeux avec la bouche on a utilisé que les distances AD, BC, SH, SM.

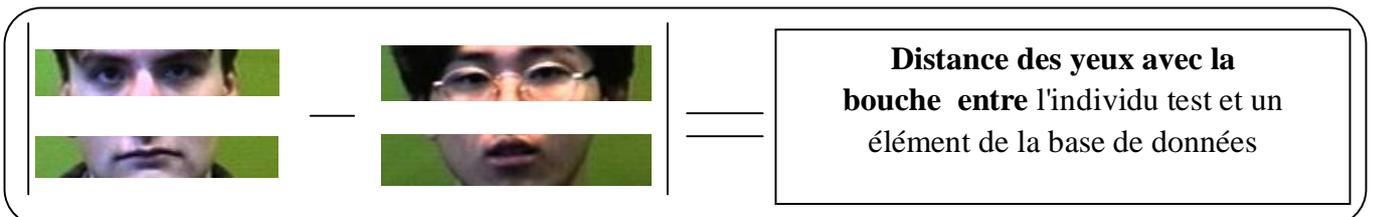


Figure 23 : Distance (yeux,bouche)

8.5. Détermination des k-ppv pour le nez avec la bouche : pour Déterminer les k-ppv pour **les yeux avec** la bouche on a utilisé que les distances EF, SG, SH, SM.

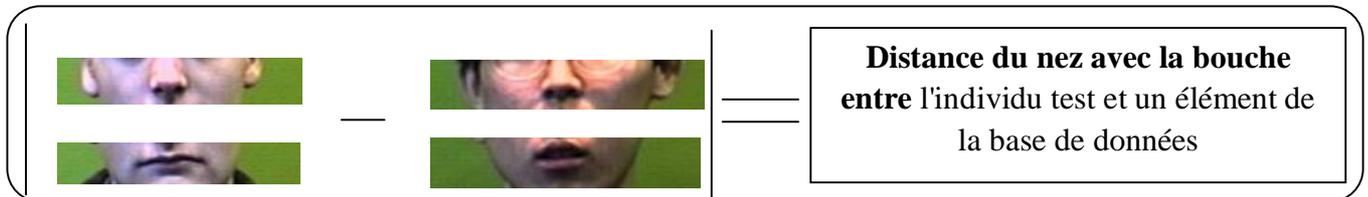


Figure 24 : Distance (nez,bouche)

9. Etude comparative :

En vue d'optimiser la qualité des résultats, nous avons choisi d'introduire de nouvelles méthodes afin d'extraire les k-ppv selon les conditions de reconnaissance. En fait, le problème major d'un système qui n'utilise que la méthode globale réside dans le fait que : si l'erreur est située dans un seul organe, ceci affecte les résultats obtenus dans les autres organes. Dans le but d'éviter ce problème nous avons choisi de segmenter le visage comme suit (yeux, nez, bouche, yeux avec nez, yeux avec bouche, nez avec bouche), cette méthode de segmentation nous permet d'obtenir de bonnes résultats.

Toutes les segmentations précédentes montrent une performance exceptionnelle dans le cas où il y a des erreurs de clique au niveau de l'étape d'extraction des caractéristiques. Pour une petite base d'image, la méthode qui traite chaque organe comme un facteur pour définir les k-ppv montre de bons résultats même si l'erreur est située dans deux organes à la fois, mais pour des grandes bases d'image la méthode qui traite la combinaison entre les organes (dichotomie) (les yeux avec le nez , bouche avec les yeux, bouche avec le nez) donne de très bons résultats dans le cas où il y a des erreurs de clique au niveau d'un seul organe. Il est à signaler que, pour des bases d'images plus grandes, il faut augmenter le nombre des points caractéristiques et utiliser la méthode globale. Cette dernière est susceptible de donner de bons résultats dans le cas où il n'y a pas des erreurs de clique.

10. Présentation de l'application :



Figure 25 : page d'accueil de logiciel



Figure 26 : Ouvrir une image de la base



Figure 27 : Menu d'apprentissage

Points
Caractéristiques

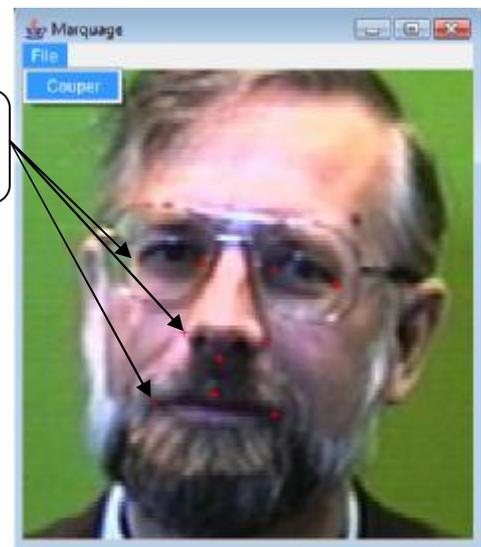


Figure 28 : Sélection des points caractéristiques pour l'apprentissage



Figure 29 : Menu de reconnaissance.

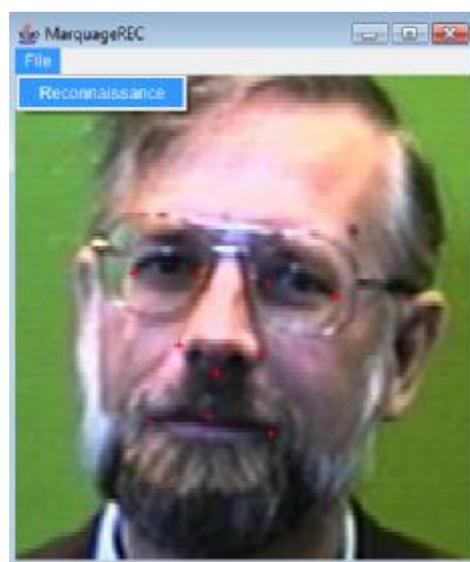


Figure 30 : Sélection des points caractéristiques pour la reconnaissance.

11. Résultats et discussions

L'évaluation de tout système de reconnaissance revient à déterminer, par un test le taux de reconnaissance, ce taux de reconnaissance est la probabilité avec laquelle une personne est identifiée, si elle existe bien sur dans la base des visages.

Notre système a été testé en utilisant la base de données Faces94, cette base contient plus de 3000 image de plus de 150 différentes personnes. Chaque individu est représenté par un ensemble de 20 images représentant les différentes expressions faciales.

$$\text{Taux de reconnaissance} = \frac{\text{Nombre d'identifications correctes}}{\text{Nombre totale des individus}}$$

	Il n'a pas des erreurs de clique	Il y a des erreurs de clique
Visage globale	97.35%	07.28%
Régions de visage	98.01%	94.70%

Tableaux 3. Taux de reconnaissance des visages.

Conclusion Générale

La méthode utilisée dans ce travail a l'intérêt d'être simple et robuste, et les résultats obtenus fournissent globalement une solution de qualité. Comparé à d'autres systèmes, ce système est caractérisé par certains avantages à savoir :

1. Taille importante de l'échantillon par rapport à d'autres systèmes. Il est encore possible d'élargir la base du système en prenant en compte un grand nombre d'individus, pour cela il faut d'abord estimer le nombre de personnes au-delà duquel le taux de reconnaissances du système va chuter considérablement.
2. Il n'y a pas de contraintes sur l'environnement comme c'est le cas pour d'autres systèmes (arrière plan neutre, pas de variation de l'éclairage...).
3. est invariant au type de coiffure, au port de chapeau, foulard, lunettes (tant que ça ne cache pas les yeux), à la présence ou absence de barbe, de moustaches et à la variation de l'expression du visage, à la couleur, ...
4. Quant à l'approche par région, si l'erreur est située dans un seul organe, contrairement aux autres approches, ceci n'affecte pas les résultats obtenus dans les autres organes.

L'exploitation de la complémentarité des mécanismes globaux et des mécanismes par région, afin d'étudier plus particulièrement les différents régions du visage, permettra au système d'optimiser le taux de reconnaissance par rapport à d'autres systèmes qu'utilise un seule mécanisme.

Ce que nous projetons dans l'avenir :

- Ø Augmenter le nombre des points caractéristiques.
- Ø Ajouter un autre classifieur comme les RNA pour comparer et améliorer les résultats.
- Ø Prendre en considération l'état émotionnel (peur, joie, tristesse, étonnement.....), ceci revient à prendre en considération la forme des sourcils et la largeur de la bouche.....
- Ø Utiliser une base plus large que la base face94.

- [01] S. Ziouche, M. Remache "Réalisation d'un système de reconnaissance des visages à base HMM", mémoire de fin d'étude EMPT 2000.
- [02] <http://www-sop.inria.fr/odyssee/team/Pierre.Kornprobst/cvision/>
- [03] K.Murat,G.Coray,G.Granlund,J.P.Haton "Reconnaissance des formes et analyse de scènes" traitement de l'information: Volume 3 Presses Polytechniques et Universitaires Romandes 2003
- [04] A.Marion"introduction aux techniques de traitement d'image", Édition Eyrolles, Année 1987.
- [05] D.ballard and browc" computer vision", Édition Prentice-hall new jersey, Année 1982.
- [06] N.Farah, "Systèmes Multiclassifieurs: Application à la reconnaissance de mots Arabes manuscrits " thèse de doctorat université d'Annaba, Département d'informatique ,LRI 2005.
- [07] J.j.toumazet " traitement de l'image sur micro-ordinateur", France, 1987.
- [08] Professeur M. VAN DROOGENBROECK, Institut Montefiore, Service de Télécommunications et d'Imagerie "Acquisition et traitement de l'image", Septembre 2001 (version 4.14).
- [09] Hastie T., Tibshirani R., Friedman J., '*The Elements of Statistical Learning.*', Springer–Verlag, 2001.
- [10] Professeur Hotuong Vinh, Etudiant : Dang hong Vu "Rapport final du TIPE, Sujet : Biométrie pour L'identification ", (Promotion X-IFI,Hanoi, Vietnam).
- [11] <http://www.biométriecgoup.com>.
- [12] http://solutions.journaldunet.com/0208/020826_biometrie_1.shtml.
- [13] f.Perronin , J.L Dugelay " Introduction à la Biométrie ;Authentification des Individus par Traitement Audio-Vidéo" Revue traitement du signal, vol19, Numéro 4,2003
- [14] B.Karl "Biométrie :Présentation Personnelle" Edition Spéciale 2002
- [15] Y.Ma, X.Ding " Face Detection based on hierarchical support vector machines",IEEE,pp.222-225,2002.
- [16] M. J Farah,. (1996). Is face recognition "special" ? Evidence from neuropsychology.Behavioral Brain Research, 76, 181-189.
-

- [17] H.Abdi, D.Valentin, "Modèles Connexionnistes et Numériques pour la mémoire des visages" *Psychologie Française*,39(6) pp 330-374 ,1999.
- [18] Tanaka, J. W. & Farah, M. J., (2003). The holistic representation of faces. Dans Peterson, M. A. & Rhodes, G. (Éds.). (2003). *Perception of faces, objects, and scenes. Analytic and holistic processes*, 53-74. Oxford: Oxford University Press.
- [19] L. Hong and A. Jain, "Integrating faces and fingerprints for personal identification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1295-1307, Dec. 1998.
- [20] Baudouin Jean-Yves . *Reconnaissance du visage, expression et genre*. Thèse de doctorat, Université Lyon II. 2002.
- [21] P.J. Phillips, P. Grother, R.J. Michaels, D.M. Blackburn, E. Tabassi, and M. Bone, "Face recognition Vendor Test 2002: Evaluation Report," NISTIR 6965, NAT. Inst. Of Standards and Technology 2003.
- [22] Laskri Mohamed Tayeb, Chefrou Djallel, *Who_is: système d'identification des visages humains*. Volume 1:pp 39-61 ARIMA. 2002.
- [23] Moses.Y : Face recognition : Generalization to novel images. Thèse de PHD, The Weizman Institute of Science, Israel, 1993.
{[Http://.FTP.wisdom.wizman.ac.il/pub/yael/th](http://FTP.wisdom.wizman.ac.il/pub/yael/th)}
- [24] T.Nagamine, Y.Uemura I.Masuda,"3D Facial image analysis for human identifiaion" *IEEE Comput .Soc .Press*,p.324-331, 1992.
- [25] H.Abdi, D.Valentin, A.J.Toole "A Generalized auto-associator model for face processing and catoration " *In d.Levine* pp 539-562,1996.
- [26] C.Faure,"Reconnaissance des formes "Juin 1990 h 1920 pp :1-15.
- [27] S.Kermi, A.kermi, c.Belfar, R.Doghmane,"Detection automatique des visages humains basée sur la couleur de la peau et les machines à vecteurs de Supports" *Proceeding of the international workshop on "Text , image & speech Recognition"* pp :150-159 ,December 12-13 2005.
- [28] Laskri Mohamed Tayeb, A.Kermi, M.A.Aouaoui, E.F. Bendris ," *Where_is: Un système d'identification des visages humains*. Mémoire de fin d'études pour l'obtention du diplôme d'ingénieur . 2002.
- [29] J.Xiaoguang ,M.S.Nixon " Extending the feature vector for automatic face Recognition" *IEEE* pp:52-60 August 1995.
-

- [30] Kamel.M.S, H.C.Shen, A.K.C. wong et R.I. Campeau :*system for the recognition of human faces*. IBM systems journal,Vol 32,No 2,pp .307-320,1993.
 - [31] R. Bruneli and T. Poggio, “Face recognition: features versus templates,” *IEEE Trans. Pattern Analysis and Machine Intelligence*,vol. 15, pp. 1042-1052, 1993.
 - [32] M.K.Hu, “Visual pattern recognition by moment invariants,” *IRE Trans. Information Theory*, vol. 8, pp. 179–187, 1982.
 - [33] D.Reisfeld, Y.Yeshurun “Robust detection of facial features by generalized symmetry” *IEEE* pp 117-120 , May 1992.
 - [34] N.Intrator, D.Reisfeld, Y.Yeshurun,” face recognition using hybrid supervised / unsupervised neural network”. *Pattern recognition Letter* 17 pp 64-76 June 1996
 - [35] S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D. Back, “Face recognition: A convolutional neural-network approach,” *IEEE Trans. Neural Networks*, vol. 8, pp. 98-113, 1997.
 - [36] S.H. Lin, S.Y. Kung, and L.J. Lin, “Face recognition/detection by probabilistic decision-based neural network,” *IEEE Trans. Neural Networks*, vol. 8, pp. 114-132, 1997.
 - [37] S.Y. Kung and J.S. Taur, “Decision-Based neural networks with signal/image classification applications,” *IEEE Trans. Neural Networks*, vol. 6, pp. 170-181, 1995.
 - [38] S.H. Lin, S.Y. Kung, and L.J. Lin, “Face recognition/detection by probabilistic decision-based neural network,” *IEEE Trans. Neural Networks*, vol. 8, pp. 114-132, 1997.
 - [39] Vapnik V.N, *Statistical Learning Theory* , Wiley, New york 1998 .
 - [40] Guo.G, Li . S.Z., and Chan .K, “Face recognition by support vector machines,” In proc. *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 196-201, 2000.
 - [41] Phillips. P.J., “Support vector machines applied to face recognition,”*Processing system* 11, 1999.
 - [42] B. Heisele, P. Ho, and T. Poggio, “Face recognition with support vector machines: Global versus component-based approach,” in *International Conference on Computer Vision (ICCV'01)*, 2001.
 - [43] K.Jonsson, J. Mates, J. Kittler and Y.P. Li, “Learning support vectors for face verification and recognition,” *Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*, pp. 208-213, Los Alamitos,
-

- USA, March 2000.
- [44] F. Samaria and F. Fallside, "Face identification and feature extraction using hidden markov models," *Image Processing: Theory and Application*, G. Vernazza, ed., Elsevier, 1993.
- [45] F. Samaria and A.C. Harter, "Parameterisation of a stochastic model for human face identification," *Proc. Second IEEE Workshop Applications of Computer Vision*, 1994.
- [46] A.S. Tolba, "A parameter-based combined classifier for invariant face recognition," *Cybernetics and Systems*, vol. 31, pp. 289-302, 2000.
- [47] John A. Black, M. Gargesha, K. Kahol, P. Kuchi, Sethuraman Panchanathan," A Framework for performance evaluation of face recognition algorithms," in *Proceedings of the International Conference on ITCOM, Internet Multimedia Systems II*, 2002.
- [48] C.Yuccer, K.Oflazer," a rotation,Scaling and translation invariant pattern classification system" *Pattern recognition Vol 6 n°5*,pp 687-710 August 1993
- [49] C.Bencheriet, «Reconnaissance de visages par approche géométrique», mémoire de magister, soutenu le 27 septembre 2000.
- [50] Seitz.P.et M. Bichsel : "The Digital Doorkeeper"- Automatic face recognition with computer. *IEEE*, pp. 77-83, 1991.
- [51] T. Kanade, "Picture processing by computer complex and recognition of human faces," technical report, Dept. Information Science, Kyoto Univ., 1973.
- [52] A.J. Goldstein, L.D. Harmon, and A.B. Lesk, "Identification of human faces," *Proc. IEEE*, vol. 59, pp. 748, 1971.
- [53] Y. Kaya and K. Kobayashi, "A basic study on human face recognition," *Frontiers of Pattern Recognition*, S. Watanabe, ed., pp. 265, 1972.
- [54] R. Bruneli and T. Poggio, "Face recognition: features versus templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*,vol. 15, pp. 1042-1052, 1993.
- [55] Vezak M, J. Korosec, L. Gyergyek, N. Pavesic, T. Savsek, I. Erjavc et A. Gere : System for description and identification of individuals. *IEEE*, Vol. 2,pp.251-1254, 1991.
- [56] M. Kirby and L. Sirovich, "Application of the Karhunen- Loève procedure for the characterisation of human faces," *IEEE Trans.Pattern Analysis and Machine Intelligence*, vol. 12, pp. 831-835, Dec.1990.
-

- [57] M. Turk and A. Pentland "Face recognition using eigenfaces," In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.586–591, 1991.
- [58] Sirovich L and M. Kirby, "Low-Dimensional procedure for the characterisation of human faces," *J. Optical Soc. of Am.*, vol. 4, pp.519-524, 1987.
- [59] T.J. Stonham, "Practical face recognition and verification with WISARD," *Aspects of Face Processing*, pp. 426-441, 1984.
- [60] Sim .Tand Kanade .T, "Combing models and exemplars for face recognition: An illuminating example," In *Proceeding Of Workshop on Models Versus Exemplars in Computer Vision*, CUPR 2001.
- [61] Z. Hammal, N.Eveno, A.Caplier, PY.Coulon''Extraction des caractéristiques faciales dans une séquence vidéo'' LIS, INPG, Rapport d'activité 2003
- [62] MAX CHASSE, Commission d'accès à l'information du Québec, Les enjeux Année 2002.
- [63] l'école Jeunes Chercheurs CNRS "INTERACTION HOMME-MACHINE", Luminy, juillet 1997
- [64] <http://www.tsi.enst.fr/~cfaure/intro/Intro1.html>
- [65] <http://www.tsi.enst.fr/~cfaure/intro/Intro2.html>
- [66] R Gilleron., M .Tommasi., 'Découverte de connaissances à partir de données.', Cours donné à l'IUP MIAGE troisième année à Lille 1, 2000
- [67] G.Saporta, '*Probabilités, Analyse des données et Statistique.*', Édition Technip, Paris, France, 1990.
- [68] L .Breiman., Friedman J., Olshen R. A., Stone C. J., '*Classification and regression trees.*', Wadsworth International Group, Belmont, CA, 1984.
-