

UNIVERSITE MOULOU D MAAMERI DE TIZI OUZOU
FACULTE DE GENIE ELECTRIQUE ET INFORMATIQUE
DEPARTEMENT INFORMATIQUE.

MEMOIRE DE FIN D'ETUDES.

DETECTION DE LA PEAU HUMAINE DANS LES
IMAGES, UTILISATION DE LA REGRESSION
LOGISTIQUE A NOYAU.

Présenté par :

MOHDEB FAWZI

&

ABROUCHE HASSAN

Soutenu devant le jury composé de :

HAMMACHE REZKI Maitre-assistant A

Président.

SADOU SAMIR Maitre-assistant A

Examineur.

DIB AHMED Maitre-assistant B

Examineur.

FILALI IDIR Maitre-assistant B

Promoteur.

UMMTO.

Octobre,2013

Nous tenons à exprimer nos remerciements et notre vive gratitude :

A Monsieur FILALI IDIR qui nous a permis, grâce à sa confiance, son aide ,et ainsi à soutien précieux, de réaliser et surtout de mener à terme ce travail.

A Monsieur HAMMACHE REZKI qui nous a fait l'honneur de présider ce jury .

Aux messieurs : SADOU SAMIR et DIB AHMED qui ont accepté D'être les examinateurs de ce mémoire, nous les remercions tout particulièrement pour l'attention et le temps qu'ils y ont consacrés. Nous remercions chaleureusement nos parents, proches, professeurs et tous ceux qui ont rendu ce travail possible par leurs conseils, soutiens, remarques et encouragements.

Nous voudrions maintenant remercier nos amis avec lesquels nous avons passé d'agréables moments pendant notre passage à l'ummta.

Et louange à ALLAH ,seigneur des univers.

REMERCIEMENTS.

TABLE DES MATIRES

TABLE DES FIGURES.

INTRODUCTION	1
A. Contexte.....	1
B. Objectifs et problématique.....	1
C. Organisation du mémoire.....	2

CHAPITRE I. **5**

APPROCHES DE DETECTION DE LA PEAU HUMAINE DANS UNE IMAGE **5**

1.1. Introduction	6
1.2. Pourquoi détecter la peau	6
1.2.1. Intérêt de détecter la peau humaine	6
1.2.2. Les domaines d'application	7
1.2.3. Quelques difficultés	7
1.3. Méthodes de détection de la peau dans une image	9
1.3.1. Détection basée sur l'extraction des traits caractéristiques	10
1.3.1.1. Détection de visages	10
1.3.1.1.1. Approches basées sur le contour	10
1.3.1.1.2. Approches basées sur les traits caractéristiques	11
1.3.1.1.3. Comparaison	11
1.3.1.2. Modélisation et suivi des mains	12
1.3.2. Approches basées sur le mouvement	12
1.3.2.1. Soustraction de l'arrière-plan (SAP) par modélisation statique	13
1.3.2.2. Différence entre deux images consécutives	14
1.3.2.3. Calcul du flot optique	15
1.3.3. Approches basées sur la couleur	15
1.4. Conclusion	15

CHAPITRE II. **18**

APPORT ET UTILISATION DES ESPACES DE COULEURS DANS LA DETECTION DE PEAU HUMAINE SUR DES IMAGES **18**

2.1. Introduction	19
2.2. Concepts	19
2.2.1. La Couleur	19
2.2.2. Les espaces de couleur	21
2.2.2.1. Les systèmes des primaires	21
2.2.2.1.1. L'espace CIE RGB de 1931	21

2.2.2.1.2. L'espace CIE XYZ	22
2.2.2.2. Les systèmes luminance-chrominance	23
2.2.2.2.1. Systèmes perceptuellement uniformes	23
2.2.2.2.1.1. (U,V,W)	23
2.2.2.2.1.2. LUV(1976)	24
2.2.2.2.1.3. CIELAB: CIELAB (L*a*b*)(1976)	24
2.2.2.2.2. Les systèmes de télévision	24
2.2.2.2.2.1. Le système (Y',I',Q')	25
2.2.2.2.2.2. Le système (Y',U',V')	25
2.2.2.2.2.3. Autres systèmes	25
2.2.2.2.3. Les systèmes antagonistes	25
2.2.2.2.4. Autres systèmes	26
2.2.2.2.4.1. Système de carron	26
2.2.2.2.4.2. CIE xyY	26
2.2.2.2.4.3. (Y,E,S)	26
2.2.2.3. Les systèmes perceptuels	27
2.2.2.3.1. HSV Le modèle de cône hexagonal (pour Hue-Saturation-Value)	27
2.2.2.4. Les systèmes d'axes indépendants	28
2.2.2.4.1. Le système d'Ohta	28
2.3. Apports des espaces de couleur pour la détection de la peau humaine	28
2.4. Conclusion	31

CHAPITRE III

32

METHODES DE DETECTION DE LA PEAU HUMAINE BASEE SUR LA COULEUR 32

3.1. Introduction	33
3.2. Fondements théoriques	33
3.2.1. Règle de décision de Bayes	34
3.2.2. Estimation paramétrique des densités de probabilité	35
3.2.2.1. Analyse discriminante avec une règle d'affectation probabiliste	36
3.2.2.2. Analyse discriminante avec une règle d'affectation géométrique	37
3.2.2.3 Estimation de la moyenne (μ) et la covariance (Σ)	39
3.2.3. Estimation non paramétrique des densités de probabilité	41
3.3. Modèle de peau paramétrique	41
3.3.1. Modèle basé sur une simple gaussienne	42
3.3.2. Modèle basé sur un mélange de Gaussiennes	42
3.3.3. Modèle elliptique de borne	43
3.4. Modèle de peau non paramétrique	44
3.4.1 Classification de pixels par table de correspondance : look up table (LUT)	44

3.4.2 Modèle basé sur l'appariement d'histogrammes	45
3.4.3 Modèle Bayésien basé sur les Histogrammes	45
3.4.4 Self-organizing map (SOM)	46
3.5. Autres modèles basée sur la couleur	47
3.6. Espaces de couleur utilisés pour la modélisation de la peau	48
3.7. Conclusion	49
CHAPITRE IV	
L'ASTUCE NOYAU (KERNEL TRICK)	50
4.1. Introduction	51
4.2. Noyaux	51
4.2.1. L'espace des caractéristiques (feature space)	51
4.2.2. Définition	52
4.2.3. Conditions pour avoir un noyau	53
4.2.4. Philosophie des méthodes à noyaux	53
4.2.5. Coût et Régularisation	54
4.2.5.1. Définition - coût régulier	54
4.2.5.2. Signification et capacité d'approximation	56
4.3. Apprentissage à base de noyaux	56
4.3.1. SVM	56
4.3.1.1. L'idée fondatrice des SVM(séparateurs à vaste marge)	56
4.3.1.2. L'hyperplan Séparateur Optimal	56
4.3.1.3. Les SVM pour le classement	57
4.3.1.4. Les SVM pour le cas séparables	57
4.3.1.5. Les C-SVM pour le cas non-séparables	58
4.3.2. Analyse en composante principale (PCA)	59
4.3.3. Processus Gaussien	60
4.4. Conclusion	61
CHAPITRE V	
CLASSIFIEUR A BASE DE REGRESSION LOGISTIQUE A NOYAU	62
5.1. Introduction	63
5.2. Régression logistique	63
5.2.1. Rappel régression logistique	63
5.2.2. Le model logit	64
5.2.3. Un cadre bayésien pour l'apprentissage supervise	64
5.2.4. Vraisemblance	64
5.2.5. Log-vraisemblance	65

5.2.6. Déviance.....	65
5.3. Régression logistique a noyau.....	65
5.3.1. Astuce de noyau (Kernel Trick).....	65
5.3.2. Régression logistique a noyau.....	65
5.4. La base d'image compaq.....	67
5.5. Performances des techniques existantes.....	68
5.5.1. Types de méthodes existantes.....	68
5.5.2. Mesure d'évaluation.....	69
5.5.3. Description du programme.....	72
5.5.3.1. Etapes.....	72
5.5.3.2. Test et résultats.....	73
5.5.4. Analyse comparative des techniques existantes.....	74
5.6. Conclusion.....	77
CHAPITRE .VI.	79
CONCLUSION S ET PERSPECTIVES	79
A. Conclusions.....	80
B. Perspectives.....	82
ANNEXES	83
Annexe A.....	83
Annexe B.....	85
BIBLIOGRAPHIE	87
Bibliographie.....	86

CHAPITRE I.

Fig 1.1. Diversité ethnique	8
Fig 1.2. Différents angles de prises	8
Fig 1.3. Présence de composantes structurelles	9
Fig 1.4. Présence de plusieurs personnes	9
Fig 1.5. Différents éclairages	9
Fig 1.6. Modèle 3D de la main	12

CHAPITRE II.

Fig 2.1. Spectre électromagnétique	20
Fig 2.2. Cube de Maxwell	22
Fig 2.3. Histogramme des pixels de peau dans l'espace rgb	29
Fig 2.4. Histogramme des pixels de peau dans l'espace YCrCb	30
Fig 2.5. Histogramme des pixels de peau dans l'espace HSI	30

CHAPITRE III.

Fig 3.1. Distribution des individus dans l'espace de description	35
Fig 3.2. Distribution relative à une loi normale	36
Fig 3.3. Frontière de séparation (Analyse discriminante avec une règle d'affectation probabiliste)	37
Fig 3.4. Frontière de séparation (Analyse discriminante avec une règle d'affectation géométrique)	38
Fig 3.5. . Frontière de séparation (Analyse discriminante avec une règle d'affectation géométrique)	39
Fig 3.6. (A gauche) un mélange de 3 gaussiennes (en pointillé) et leurs densités théoriques (trait plein) (A droite) l'histogramme associé	43
Fig 3.7. Architecture possibles de la carte SOM : rectangulaire (à gauche) et hexagonale (à droite)	47
Fig 3.8. Base de peaux de Von Luschan	48

CHAPITRE IV.

Fig 4.1. Cas non linéairement séparable	52
Fig 4.2. Chaîne de traitements générique des méthodes à noyaux	54
Tab 4.1. Récapitulatif de divers algorithmes d'apprentissage selon la nature de leurs fonctions objectifs	55
Fig 4.3. Exemples de séparateurs linéaires minimisant le risque empirique (cas séparable)	57
Fig 4.4. L'hyperplan séparateur dans le cas de données linéairement séparables	58

CHAPITRE V.

Fig. 5.1: Quelques images du corpus Compaq.....	68
Tab.5.1: Matrice de contingence bi classe.....	69
Tab. 5.2: Tous les pixels sont de la classe peau.....	71
Tab. 5.3: Tous les pixels sont de la classe non peau.....	71
Tab. 5.4: Classifieur parfait.....	71
Tab. 5.5: Classifieur le pire.....	71
Fig. 5.2. Déroulement du programme.....	73
Tab 5.6.Résultats sur différents espaces de couleurs.....	74
Tab. 5.7: Comparaison des différentes méthodes de détection de la peau.....	75

INTRODUCTION :

A. Contexte

L'augmentation régulière de la puissance de calcul des microprocesseurs, les progrès réalisés dans les méthodes de développement des logiciels, la numérisation du son, puis des images (i.e., le multimédia), la compression des signaux et le déploiement des réseaux accélérés par les fibres optiques et les satellites dessinent de plus en plus clairement les contours d'un nouveau paysage : la Société de l'Information. Cette société est construite autour de voies électroniques ou de réseaux chargés d'acheminer dans les entreprises, les universités, les administrations, les écoles et les maisons une palette très large de services interactifs. Parmi ces derniers, on peut citer la messagerie, la visiophonie, la télé-enseignement, la consultation de banques de données, la télé-achat, la télévision à la demande, etc. bouleversant de manière radicale et irréversible la vie des individus et des institutions.

B. Problématique et Objectifs

Une information de plus en plus visuelle est une conséquence majeure de cette convergence entre l'informatique, Internet et l'audiovisuel. De plus en plus d'applications produisent, utilisent et diffusent des données visuelles, incluant des images fixes et animées.

L'augmentation significative des informations visuelles au sein de Internet et dans les organisations s'est accompagnée d'une prise de conscience de l'importance de développer des moyens informatiques pour traiter ces informations. Ce traitement permet de modéliser, de filtrer, de classer, de rechercher et d'indexer cette quantité importante de données. Si la numérisation d'image est déjà un problème techniquement résolu, il n'est pas de même avec la classification, le catalogage et l'indexation d'image. Dans un tel contexte, il est impératif de pouvoir classer les images selon leurs thèmes ou leurs contenus. Cette classification permettra de faire une sélection ou un contrôle d'accès selon la sémantique et selon le type des images. Ce travail s'effectue actuellement de manière manuelle par annotation, ce qui présente un coût en temps et en main d'oeuvre trop important. De plus, la taille des collections d'images est de plus en plus gigantesque, ce qui rend leur annotation manuelle quasi impossible. En conséquence, nous nous trouvons actuellement devant une masse d'informations visuelles mal classifiée et incontrôlée sur Internet.

C'est précisément dans ce contexte que s'inscrit le travail que nous avons réalisé dans le cadre de ce mémoire de fin d'études. Nous nous sommes principalement intéressés au problème d'identification de pixels de peau dans l'image. Ce sujet de recherche est d'enjeu important

dans la mesure où il est indispensable avant d'envisager des analyses et des traitements de niveau supérieur.

Nous allons réaliser une discrimination sur un corpus d'images (Compaq). Et ce en deux étapes principales : entraînement et test et ce avec le concours d'un modèle d'apprentissage statistique qui la régression logistique à noyau.

Dans la littérature, il existe de nombreuses méthodes pour détecter dans l'image les régions de peau, comme le visage et/ou les mains d'une personne. Trois approches principales se détachent : l'une, basée sur la couleur, l'autre, basée sur l'extraction de traits caractéristiques et la dernière basée sur le mouvement. Notre objectif est de développer un modèle de peau permettant de détecter toutes les régions de peau d'une image en un temps raisonnable. De ce fait, nous avons choisi une méthode basée sur la détection de peau par une approche couleur. De nombreux travaux ont été réalisés sur l'élaboration d'un modèle de peau basé sur l'indice couleur. La difficulté réside dans la prise en compte des conditions de lumière, de la richesse ethnique avec des teintes variées et des décors complexes. En effet, la plupart des travaux pratiquent une phase d'apprentissage sur des classes prédéfinies d'images, sous des conditions d'éclairage connues à l'avance. Si ces modèles conviennent généralement aux systèmes à base d'images peu variées, ils sont peu adaptés aux systèmes contenant une grande variété d'images.

C. Organisation du mémoire :

L'organisation du présent manuscrit est présentée ci-dessous.

Le Chapitre I porte sur la présentation des différentes approches de détection de la peau humaine (Nous en avons abordé 3).

Le chapitre II est consacré à la colorimétrie et aux espace de couleur . nous avons présenté un certain nombre d'espaces et de sous espace.

Le chapitre III est consacré à l'état de l'art sur les travaux d'identification des pixels de peau dans l'image basés sur l'approche couleur . Les différentes méthodes ont été classées et présentées. Leurs avantages et défauts sont discutés.

Le chapitre IV évoque l'astuce noyau (kernel trick) et certains algorithmes qui y sont associés :les svm , le PCA et le processus gaussien.

Le chapitre V est consacré à notre application .entamé avec la présentation de la régression logistique .suivie de la présentation de l'technique sur laquelle est basé notre travail en

l'occurrence la régression logistique à noyau. Nous avons aussi présenté la base d'image Compaq de laquelle on a pris nous échantillons d'images pour l'apprentissage et le test.

CHAPITRE .I.

Approches de détection de la peau humaine dans une image couleur.

1.1.Introduction :

De nos jours, la détection des régions de peau dans l'image est devenue un sujet de recherche et un enjeu important dans la mesure où elle est indispensable avant d'envisager d'autres analyses et traitements sur les images telles que la détection et la reconnaissance de visages ou encore la classification d'images en image adulte/non-adulte. En effet, de nombreux systèmes de détection sont sujets de recherche très important étant donné qu'ils ont une utilité potentielle pour de nombreuses applications telles que les applications liées à la sécurité (surveillance, droit d'accès, ...). Ces systèmes peuvent être utilisés aussi pour faciliter la recherche et la navigation dans une masse de vidéos tel que l'identification de séquences vidéo où un visage, une partie du corps particulière apparaît. La mise en place d'un tel système de détection et reconnaissance de visages doit s'appuyer en préalable sur un système de détection et segmentation de régions de peau.

Un exemple d'envergure, Internet apparaît comme un immense gisement d'informations. Son libre accès conduit à des usages indésirables comme par exemple l'accès par des enfants à des sites adultes, il est donc nécessaire d'introduire des outils de filtrage de sites pour des applications comme le contrôle parental. Ces outils doivent s'appuyer sur une analyse sémantique d'images dans le processus d'identification, où, le texte à lui seul n'est plus suffisant. La détection des régions de peau dans l'image et le pourcentage de ces derniers sont des indicateurs pertinents pour classer les images en image adulte/non-adulte.

Dans ce chapitre nous allons aborder, avec des définitions et exemple la méthode de détection basée sur la peau humaine et les approches utilisées pour y parvenir.

1.2. Pourquoi détecter la peau :

1.2.1. Intérêt de détecter la peau humaine :

De nombreux systèmes de vision ont pour but la détection, la localisation et la reconnaissance de personnes. Par exemple, les systèmes de vidéo-surveillance font tous appel à des méthodes de détection de personnes.

La détection basée sur l'analyse de la couleur de la peau, est un moyen efficace car, d'une manière générale, au moins quelques parties du corps telles que la tête et le visage, les bras et les mains sont visibles, Les méthodes de détection basées sur l'analyse de la couleur de la peau sont des méthodes efficaces et rapides. Elles réduisent l'espace de recherche des régions concernées dans l'image.

De plus, la couleur de la peau est une information robuste face aux rotations, aux changements d'échelle, et aux occultations partielles.

Cependant, la détection de la peau n'est pas une tâche facile dans la mesure où la couleur ou la teinte de la peau, comme la couleur de tout autre objet, varie en fonction du matériel utilisé pour acquérir l'image ainsi que des conditions d'acquisition, i.e. l'environnement dans lequel l'image est acquise. De plus, dans des scènes complexes telles que les scènes d'extérieur, il peut être difficile, voire impossible, de distinguer la peau d'une personne de tout autre objet présent sur cette image.

Dans ce premier chapitre nous allons présenter les différentes approches utilisées pour la détection de la peau dans une image donnée.

1.2.2. Les domaines d'application :

La détection de la peau dans une image, est une étape importante dans de nombreuses applications de la vision par ordinateur, dans plusieurs domaines et elle est une étape indispensable dans le traitement de certains problèmes à l'image :

- la détection et reconnaissance de visages. [1] [2] [3] ;
- la détection des mains ou autre organe apparent dans le suivi d'objets sur séquences d'images. [4] [5];
- détection de corps humains nus dans une image. [6] [7].

La résolution partielle de ces problèmes conduit à réaliser des progrès dans les applications suivantes :

- la recherche des images sur le web. [8] [9] ;
- le filtrage du web, afin, par exemple, de signaler les images adultes.[10]
- l'indexation des images par le contenu.
- la segmentation de la vidéo. [11] ;
- Dans les applications intégrant des algorithmes de détection de la peau, l'évaluation de la performance est basée sur l'efficacité de tout le système, hormis dans des applications dédiées au traitement de problèmes nécessitant uniquement cette technique.

1.2.3. Quelques Difficultés :

Le processus de détection de la peau dans une application donnée, peut se heurter à de nombreuses entraves. La majorité des solutions proposées jusqu'à maintenant sont loin d'être satisfaisantes car elles fonctionnent uniquement dans des environnements contrôlés, et par

CH.I. Approches de détection de la peau humaine dans des images.

conséquent elles ne gèrent pas la variabilité des conditions d'acquisition de la vie quotidienne, notamment :

- **Diversité ethnique** : teintes qui varient selon la personne et selon les différentes races. (Fig1.2).
- **La pose** : où les images d'un visage changent en fonction de l'orientation de ce dernier (frontal, 45 degrés, profil). (Fig.1.3)
- **La présence ou absence des composantes structurales** : les caractéristiques faciales tels que la barbe, la moustache, les tatouages et les lunettes causent une grande variabilité des composantes structurales des régions de la peau, notamment au niveau de la forme, de la couleur, et de la taille. (Fig.1.4)
- **Les occultations** : les visages peuvent être partiellement occultés par d'autres objets. En effet, dans une image contenant un groupe de personnes par exemple, des visages peuvent partiellement masquer d'autres visages.(Fig.1.5)
- **Les conditions d'illumination** : des facteurs tels que l'éclairage (distribution de la source de lumière, son intensité, son spectre) et les caractéristiques de l'appareil photographique affectent l'aspect d'un visage dans l'image acquise. (Figure 1.6)



Fig.1.1. Diversité ethnique.



Fig.1.2. Différents angles de prises.



Fig.1.3. Présence de composantes structurelles



Fig1.4. présence de plusieurs personnes.



Fig.1.5. Différents éclairages.

La détection de peau est encore plus difficile pour les images issues du Web qui sont capturées sous diverses conditions de lumière et avec différents dispositifs ayant des caractéristiques spécifiques.

1.3. Méthodes de détection de la peau dans une image :

Il existe de nombreuses méthodes pour détecter les régions de peau dans une image. On en identifie trois approches majeures que nous développons dans la suite :

- la détection basée sur l'extraction de traits caractéristiques du visage et /ou des mains [12],[13]

- La détection basée sur la couleur.[14],[1]
- la détection basée sur le mouvement. Cette dernière n'est jamais utilisée seule pour la détection, on la trouve utilisée conjointement avec l'information de couleur de peau et couplée avec un système basé sur la reconnaissance de traits caractéristiques [15],[16].

1.3.1. Détection basée sur l'extraction des traits caractéristiques :

Cette méthode est généralement utilisée pour la segmentation des régions de peau des visages / ou des mains, et ce dans le but de reconnaître des personnes ou des gestes, Une telle segmentation nécessite généralement la définition de modèles pour les traits caractéristiques du modèle cherché, et parfois même un modèle du corps humain plus ou moins détaillé.

1.3.1.1. Détection de visages :

On recherche souvent des traits caractéristiques tels que les yeux, les contours extérieurs, le nez et la bouche que l'on associe à des configurations ("templates") connues a priori ou apprises [17].

Les contraintes géométriques, permettent de définir le positionnement relatif de ces traits caractéristiques et d'en obtenir une information sur la présence ou non d'un visage.

1.3.1.1.1. Approches basées sur le contour :

a/ le système proposé par Govindaraju et al. [18] :

Il modélise le visage comme étant un agencement de trois courbes (sommet, coté droit et coté gauche du visage). La détection de ces courbes et leur regroupement selon leur position relative permet de connaître la position du visage.

b/ Une autre approche basée sur les contours consiste en une extraction des contours de l'image et leur mise en adéquation avec une ellipse modélisant un visage [19] [20].

L'utilisation de contours actifs permet une adaptation aux variations de forme du visage et conduit ainsi à une meilleure délimitation du visage.

les méthodes basées sur le contour les plus abouties reposent sur l'utilisation de contours actifs .exemple : Waite *et al.* [21], Craw *et al.* [22] et Cootes *et al.* [23].

1.3.1.1.2. Approches basées sur les traits caractéristiques :

Il existe une multitude de méthodes exploitant les traits caractéristiques, et ce est principalement du au nombre important de caractéristiques définissables et des techniques de détection.

Les traits du visage les plus souvent retenus sont les yeux, les sourcils, la bouche et le nez, qui peuvent être d'un niveau de détail beaucoup plus fin. La détection est alors basée sur le fait que les caractéristiques d'un visage ont des positions relatives fixes (ou statistiquement modélisable).

a/ La méthode de Kanade en 1973 [24] [25] : constitue un premier pas vers les techniques de détection et de reconnaissance de visage.

Principe :

- Extraction des contours de l'image
- Projection des contours selon des axes horizontaux et verticaux.

Selon le modèle obtenu, le système est capable de retrouver le positionnement des yeux, de la bouche, du nez, ainsi que les contours du visage. Si le profil obtenu ne correspond pas au modèle de référence, l'image est rejetée comme n'étant pas un visage. La méthode la plus répandue pour extraire ces caractéristiques est de réaliser une corrélation entre l'image et un masque de caractéristiques génériques (Sumi et Otha [26], Zelinsky et Heinzmann [27]).

*b/ Une autre méthode utilise des filtres successifs pour l'extraction des paramètres du visage, ce qui constitue une approche généralisée de la méthode précédente (Leung *et al.*[28], Graf *et al.* [29], Burl *et al.* [30]).*

Inconvénient : l'ensemble des caractéristiques retenues ne constitue pas un modèle robuste et conduit à un grand nombre de fausses alertes.

Solution : utilisation de techniques de comparaison avec un modèle de déformation, alors la robustesse du système est grandement améliorée [20].

c/ Dans [15], Yow et Cipolla décrivent un système de détection du visage où les yeux, la bouche et le nez sont extraits par un filtrage en quadrature de phase et sont regroupés selon leur probabilité à constituer un visage. Le problème de l'angle de vue est résolu par l'utilisation d'invariants augmentant la robustesse de leur approche.

1.3.1.1.3. Comparaison :

Les approches basées sur l'extraction de traits du visage sont plus efficaces que les approches basées sur les contours car elles analysent des structures locales de l'image conjointement à

des modèles géométriques de visages. De plus, il s'agit d'approches ascendantes permettant d'écarter les fausses alertes à chaque étape de la détection d'un nouveau trait du visage [20].

1.3.1.2. Modélisation et suivi des mains :

Certains chercheurs fournissent une forme 2D de la main. D'autres tels que [31] préfèrent utiliser un modèle générique 3D, où la main est représentée par un modèle de type volumique articulé (cf. figure 2.5). Celui-ci, conformément au paramétrage MPEG-4 [32], peut être animé suivant les 6 degrés de liberté pour le positionnement global, et 5 degrés de liberté pour chaque doigt. Le modèle générique 3D ne correspond pas en général à la morphologie particulière de la main d'un opérateur. L'ajustement préalable des paramètres de morphologie du modèle est nécessaire pour la robustesse de la procédure de suivi de la main dans des séquences. Cette opération est effectuée à partir d'une seule image de la main ouverte (doigts écartés). La silhouette de la main est obtenue à partir d'un seuillage sur la teinte de la peau et les doigts sont séparés par filtrage morphologique. Les paramètres de la morphologie des doigts (rayon et longueur) et de la paume (longueur et largeur) sont calculés à partir des dimensions des régions correspondantes [33].

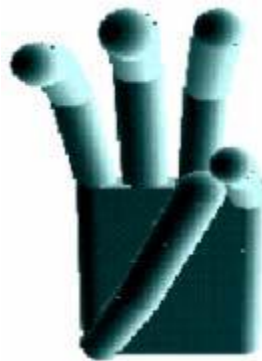


Fig.1.7 .Modèle 3D de la main.

1.3.2. Approches basées sur le mouvement :

L'utilisation de l'information de mouvement peut être un moyen simple pour mettre en œuvre une technique rapide de détection de peau dans une vidéo.

Principe : ce système suppose généralement que l'arrière plan de la scène vidéo est stationnaire et que les régions contenant la peau tels que le visage ou/et les mains par exemple sont en mouvement. Dans ce cas, ces régions peuvent être détectées par une simple différence entre l'image courante et l'image précédente.

L'information de mouvement est souvent combinée avec l'information de couleur de peau ou couplée à un système basé sur la reconnaissance de traits caractéristiques.

L'exploitation de l'information de mouvement oriente la détection de peau vers des zones préférentielles en éliminant les zones sans intérêt et permet donc une forte réduction de la complexité de la détection et par conséquent une limitation importante des calculs et des traitements [34].

Principal inconvénient : La présence de plusieurs objets en mouvement sur la vidéo réduit considérablement les performances de ces systèmes.

Les allons présenter quelques techniques basées sur ce principe.

1.3.2.1. Soustraction de l'arrière plan (SAP) par modélisation statique ([35], [36] et [37]) :

C'est la plus grande famille des méthode de détection de mouvement est ce est du probablement à :

- Simplicité théorique.
- Faible complexité algorithmique.

Principe : estimation statistique de la scène observée. Le mouvement est détecté en comparant une image test avec le modèle d'arrière-plan calculé auparavant. Certaines hypothèses de base doivent par contre être respectées pour un fonctionnement adéquat de cette méthode.

Exemple : la caméra utilisée est fixe et ne doit bouger à aucun moment. Une caméra à l'épaule est un bon exemple de situation où la SAP ne peut pas s'appliquer.

Il est nécessaire de mettre à jour le modèle statique calculé et ce dans le but s'adapter aux changements qui se produisent sur la scène observée. Cette capacité d'adaptation est commune à toutes les techniques de SAP par modélisation statistique et leur confère un atout majeur. Par ailleurs, cette méthode connaît plusieurs implantations différentes qui varient principalement selon le type de capteur utilisé.

a/ Visible (2D) :

Principe : utilisation d'images 2D dans le spectre visible.

Le processus se déroule comme suit :

- modéliser l'arrière-plan à partir de plusieurs images acquises séquentiellement.
- Pour chaque pixel de l'image, ainsi que pour chacun des canaux (R, G et B), une moyenne et une variance sont calculées.

- Lorsqu'un pixel test doit être classifié, il faut tout d'abord lui soustraire la moyenne correspondante dans le modèle statistique. Il sera alors étiqueté comme un pixel contenant du mouvement seulement si la valeur absolue du résultat dépasse un certain multiple de l'écart type correspondant.

d'autres modèles de couleur ont été utilisés auparavant pour la modélisation statistique, comme par exemple le YUV par Wren *et al.* [38] et le HSV par Cucchiara *et al.* [39].

b/ Horprasert et al. [35] ont proposé un nouveau modèle de couleur basé sur le RGB.

Principe : classification des pixels en quatre catégories : appartenant à arrière plan original, illuminé, ombré, et en mouvement. Pour cela, deux mesures sont ajoutées à la méthode de base en RGB, qui est :

La distorsion chromatique (α) distorsion de luminosité (CD).

Inconvénients :

- somme d'opérations supplémentaires nécessaires pour calculer ces deux mesures
- les seuils associés.

En pratique, certaines erreurs de classification peuvent également se produire entraînant, par exemple, l'identification d'un objet en mouvement comme étant de l'ombre .

1.3.2.2. Différence entre deux images consécutives :

La différence entre deux images consécutives représente une solution très intéressante.

Principe : elle consiste à soustraire une image acquise au temps t_n d'une autre au temps t_n+k , où k est habituellement égal à 1.

Résultats :

- image vide si aucun mouvement ne s'est produit pendant l'intervalle de temps observé car l'intensité et la couleur des pixels seront presque identiques.
- Par contre, si du mouvement a lieu dans le champ de vue, les pixels frontières des objets en déplacement devraient changer drastiquement de valeurs, révélant alors la présence d'activité dans la scène.

Cette technique étant peu complexe et nécessite très peu de ressources, car aucun modèle n'est nécessaire. Cela implique donc qu'il n'y a pas de phase d'initialisation obligatoire avec une scène statique, ce qui procure une très grande flexibilité d'utilisation. De plus, une opération de soustraction d'images requiert très peu de puissance de calcul, lui conférant un avantage supplémentaire [40].

Inconvénients : Cependant, les résultats obtenus avec cette méthode ne sont pas aussi intéressants que ceux générés en utilisant un modèle statistique de l'arrière plan. En effet, certains traitements supplémentaires sont nécessaires afin de déterminer la zone en mouvement, car l'information disponible ne concerne que les contours des régions en déplacement (ce qui inclus également les zones intérieures d'un objet).

1.3.2.3. Calcul du flot optique

Les méthodes de calcul de flot optique sont basées sur deux hypothèses:

- la photométrie constante d'un pixel
- le faible espace temporel entre les images sur lesquelles sont calculées le flot optique.

En prenant des images très proches les unes des autres, (dans une séquence vidéo par exemple, les images sont prises entre 1/10 et 1/25 seconde d'écart), le mouvement est faible et les pixels en mouvement gardent presque les mêmes valeurs d'intensité.

Le principe consiste à résoudre, pour un pixel ou un bloc, l'équation du flot optique, appelée aussi l'équation de constance de luminance ou l'équation de contrainte du mouvement apparent (ECMA). Elle est la base de toutes les méthodes de flot optique.

$$uI_x + vI_y + I_t = 0 \quad (1.1)$$

avec
$$u = \frac{\Delta x}{\Delta t}, v = \frac{\Delta y}{\Delta t}, I_t = \frac{\Delta I}{\Delta t} \quad (1.2)$$

1.3.3. Approches basées sur la couleur :

La détection de la couleur de la peau est souvent utilisé pour définir une série de zones candidates. De plus, celle-ci peut être effectuée de manière assez simple grâce à des seuillages dans un espace de couleurs approprié. Une telle détection est alors relativement peu coûteuse en temps de calcul et donc particulièrement adaptée à un détecteur devant traiter un nombre significatif d'images par seconde. L'objectif donc est la classification automatique de pixels en pixels de peau ou non-peau dans les images.

De nombreux chercheurs tels que, Fleck *et al.* [6] [7], Kjeldsen *et al.* [40], ont montré que la couleur de la peau est localisée dans une bande étroite de l'espace de couleur. Cette information peut donc facilement être utilisée pour marquer les pixels de couleur peau.

On distingue principalement deux axes de recherche dans cette approche couleur.

CH.I. Approches de détection de la peau humaine dans des images.

- Modélisation de peau/non peau permettant de distinguer les pixels de couleur de peau de ceux de non-peau.
- Choix de l'espace de couleur que nous devons adopter pour une bonne classification.

Les méthodes de classification se déclinent généralement en 2 familles :

- le mode supervisé
- mode non supervisé.

On peut distinguer également deux catégories de méthodes de classification :

- les méthodes indirectes qui utilisent la formule de Bayes.
- les méthodes directes qui évaluent les probabilités *a posteriori* sans faire intervenir la formule de Bayes.

La formule de Bayes permet de déterminer les probabilités d'appartenance *a posteriori* si les densités de probabilité et les probabilités a priori sont connues. La règle de Bayes permet d'obtenir le taux d'erreur de classification minimum, ce qui est l'objectif souhaitable pour tout système de classification.

Cette approche fera l'objet du troisième chapitre, Nous y reviendrons en détails tout au long du troisième chapitre.

1.4. Conclusion :

Dans ce chapitre nous avons pu présenter d'une manière plus ou moins détaillée certaines méthodes pour la détection de régions de peau dans une image. Trois approches se détachent principalement : la détection basée sur l'extraction de traits caractéristiques du visage et /ou des mains, la détection basée sur une approche couleur, et la détection basée sur le mouvement. Cette dernière n'est jamais utilisée seule.

Les techniques basées sur l'extraction de traits caractéristiques nécessitent la définition de modèles pour ces traits (main, visage, corps humain...) qui sont associés à des templates. Le template matching consiste à comparer l'intensité de pixels entre un template prédéfini et plusieurs sous régions de l'image que l'on désire analyser.. Cette approche souffre de deux handicaps : la variation de formes d'un objet et une connaissance supposée de l'échelle de l'objet.

Les techniques basées sur le mouvement, telle que la méthode de différence d'image sont des techniques simples permettant de faire rapidement une estimation de la position d'un objet en

CH.I. Approches de détection de la peau humaine dans des images.

mouvement. Cette estimation permet de réduire la zone de recherche d'un autre algorithme. On trouve cette approche souvent utilisée conjointement avec l'information de couleur de peau ou couplée avec un système basé sur la reconnaissance de traits caractéristiques.

CHAPITRE II.

*Apport et utilisation des espaces de couleurs dans la
détection de peau humaine sur des images.*

2.1. Introduction:

Dans le domaine de la couleur, plusieurs représentation électroniques coexistent, avec trois ou quatre valeurs de base selon le cas, à l'exemple de (RVB),(LAB),(HSV),(HLS),(HIS),(CMYN) etc ...

Ces représentations sont le plus souvent imposées par telle ou telle technologie, ce qui conduit d'ailleurs à les multiplier.

Ainsi, Le (RGB)/(RVB), des ordinateurs et des tubes de télévision, qui n'est pas bien adapté à la transmission hertzienne, se voit remplacé par (YUV, luminance et deux chrominance) pour les standards de télévision européenne, ou (YIQ) aux USA.[41]

En outre, la perception humaine est moins sensible aux modes cartésiens comme RVB ou YUV qu'aux représentations en coordonnées polaires, de type brillance, saturation teinte, ces derniers d'ailleurs sont proposées pour effectuer les réglages. Finalement, dans les postes de télévision grand public, on utilise trois représentations à divers moments de la chaîne de traitement.

Par contre lorsque nous nous proposons d'effectuer des traitements sur les images électronique de couleur, de telles contraintes technologiques n'ont plus aucune raison d'être .on peut soit garder le triplet RGB ,qui un standard informatique ou bien le remplacer par tout autre système qui lui correspond de manière réversible, dès que cet autre système est mieux adapté à ce que l'on veut faire.

2.2. Concepts :

2.2.1. La Couleur :

On nomme **couleur** la perception par l'œil d'une ou plusieurs fréquences d'ondes lumineuses, avec une (ou des) amplitude(s) donné(s) [42]. Il importe de ne jamais confondre **couleur**, notion perceptive, et longueur d'onde, notion physique. Ainsi, l'œil humain est le plus souvent incapable de distinguer un jaune monochromatique (une seule longueur d'onde) d'une composition correspondante de vert et de rouge. Cette illusion permet d'afficher du jaune sur nos écrans d'ordinateur, et, plus généralement, la synthèse trichrome [43].

Lorsque les images en entrée sont en couleurs, il est avantageux d'utiliser cette information supplémentaire pour isoler les régions susceptibles de contenir les objets qu'on veut identifier.

Dans le cas de la détection de la peau, plusieurs travaux ont été réalisés. Dans la majorité des cas, la peau est représentée par une portion d'un espace de couleurs particulier. En utilisant les frontières de cette région comme valeurs de seuillage sur une image, il est possible d'extraire les pixels dont la couleur peut s'apparenter à celle de la peau.

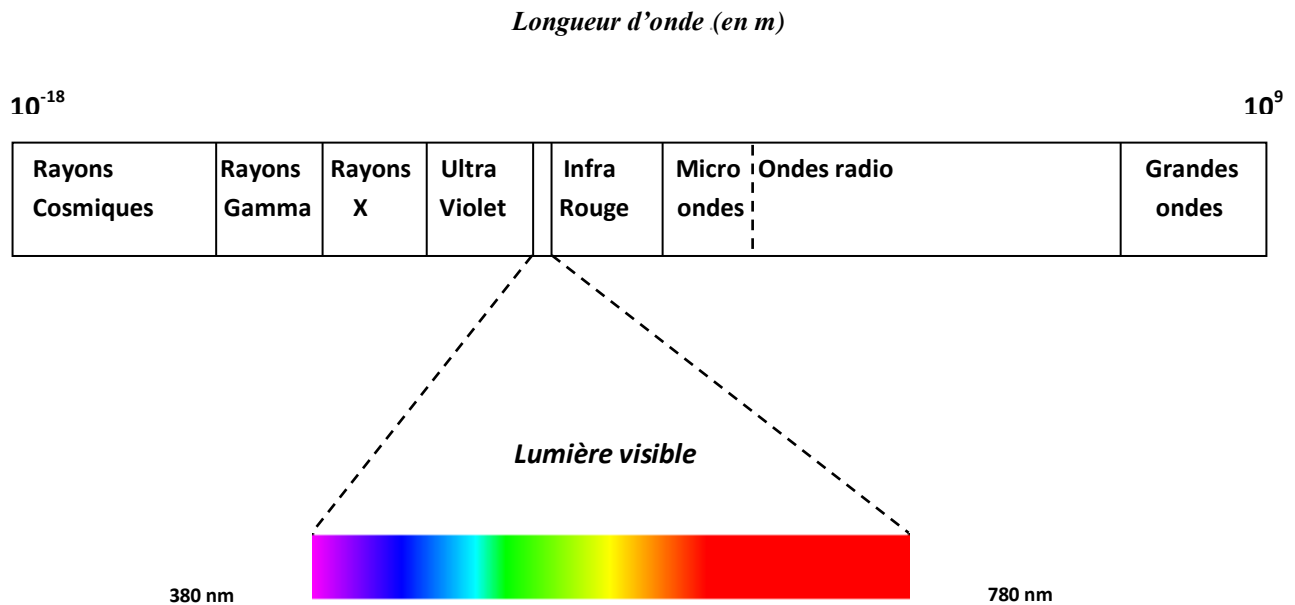


Fig.2.1.Spectre électromagnétique.

Il existe de nombreux systèmes de représentation de la couleur ainsi que de nombreuses définitions pour un même système.

Les premiers systèmes proposés pour manipuler la couleur furent basés sur des impressions psychosensorielles afin d'élaborer des classifications de couleurs. Ce sont donc des systèmes ordonnés de couleurs qui sont des représentations géométriques visuelles sous forme de planches, de nuanciers, de collections ou d'atlas [44].

Ces représentations sont en majorité réalisées dans des espaces à trois dimensions et les unités des composantes couleur sont purement conventionnelles car elles ne servent qu'à repérer une couleur à l'intérieur d'un solide dans lequel elle est inscrite. Ces systèmes comme le système Munsell, le système NCS (Natural Color System), le système OSA (Optical Society of America) ou le système DIN (Deutsches Institut für Normung) dont l'usage est le plus répandu sont peu utilisés dans le domaine de l'analyse d'images couleur[45].

Les systèmes utilisés en analyse d'images couleur sont des systèmes dans lesquels chaque couleur est définie numériquement de sorte de disposer d'une métrique permettant d'évaluer leur similarité.

Dans ce premier chapitre nous allons de faire une synthèse des systèmes de représentation de la couleur les plus utilisés, avec de brèves définitions.

2.2.2. Les espaces de couleur :

Les espaces couleur peuvent être classés suivant quatre grandes familles : les systèmes de primaires, les systèmes perceptuellement uniformes, les systèmes perceptuels, qui sont basés sur

CH II. Apport et utilisation des espaces de couleur dans la détection de la peau humaine sur des images.

les attributs de perception de la couleur, et enfin les systèmes d'axes indépendants. Définis à l'origine dans le cadre strict de la colorimétrie, ils sont à présent largement exploités en vision par ordinateur. Dans cette section, nous soulignons également l'apport de ces espaces dans le domaine de détection de la peau humaine. [46]

2.2.2.1. Les systèmes des primaires :

2.2.2.1.1. L'espace CIE RGB de 1931 :

Le seul système que nous pouvons appeler (R,G,B) est sans doute celui défini par la CIE [47]. Cet espace est défini à partir de trois primaires monochromatiques de couleurs rouge, vertes et bleu. Les longueurs d'ondes associées à chacune des primaires sont les suivantes :

- 700,0 nanomètres pour le rouge,
- 546,1 pour le vert,
- 435,8 pour le bleu

De plus la puissance de chacune de ses primaires est ajustée de façon à obtenir trois triplets identiques pour tout spectre d'égale énergie.

La CIE a défini en 1964 un autre espace de couleur RGB dont les fonctions d'appariement ont été obtenues avec un écart de 10 degré (donc pour des taches colorées plus éloignées).

il faut néanmoins noter qu'il existe de nombreux espaces RGB qui se différencient entre eux par l'illuminant utilisé mais aussi par le nombre de bits de codage que requiert l'application (espaces Y IQ, Y UV et Y C1C2). Ces systèmes sont utilisés par exemple en télévision. [48]

La représentation des couleurs dans cet espace donne un cube appelé *cube de Maxwell*.

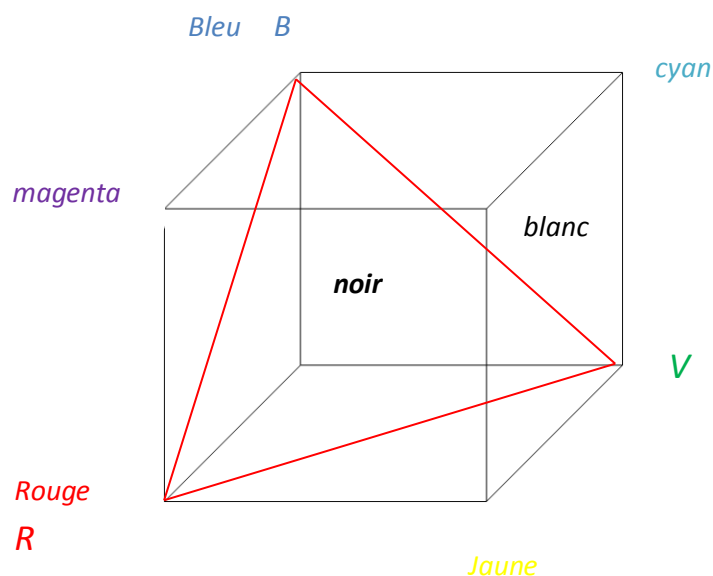


Fig.2.2. Cube de Maxwell.

CH II. Apport et utilisation des espaces de couleur dans la détection de la peau humaine sur des images.

Les composantes trichromatiques d'un stimulus de couleur sont liées à sa luminance. Deux stimuli de couleur peuvent ainsi posséder le même caractère chromatique, que nous appellerons *chrominance*, mais avoir des composantes trichromatiques R_c , B_c et G_c différentes à cause de leur luminance. Afin d'obtenir des composantes qui ne tiennent compte que de la chrominance, il convient donc de normaliser les valeurs des composantes trichromatiques par rapport à la luminance. Ceci est réalisé en divisant chaque composante trichromatique par la somme des trois. Les composantes ainsi obtenues sont appelées [49] :

Coordonnées trichromatiques, coordonnées réduites ou encore composantes normalisées.

Elles sont notées r_c, b_c et g_c et définies par :

$$r_c = \frac{R_c}{R_c+B_c+G_c} \quad , \quad g_c = \frac{G_c}{R_c+B_c+G_c} \quad , \quad b_c = \frac{B_c}{R_c+B_c+G_c} \quad (2.1)$$

2.2.2.1.2. L'espace CIE XYZ :

Si nous utilisons trois spectres visibles comme base de l'espace visuel humain, il existera toujours des spectres ayant des coordonnées négatives sur cette base.

Si nous utilisons par exemple l'espace RGB ou les trois primaires sont visibles, les couleurs bleu-vert(450-550 nm) ne peuvent être reproduites par superposition des trois spectres. Ceci peut poser des problèmes dès que l'on désire travailler en synthèse additive (en n'ajoutant que des quantités positives de couleur). La CIE a donc conçu un espace de couleur basé sur trois primaires X, Y et Z non visibles. Cet espace possède plusieurs propriétés intéressantes[49] :

- Les triplets décrivant chaque couleur en fonction de ses primaires ont tous des valeurs positives pour les spectres visibles.
- Une fonction Y représentant approximativement la sensibilité de l'œil humain à la luminosité. La composante Y du triplet (X,Y,Z) est usuellement assimilée à la luminance du spectre incident.
- Tout spectre d'égale énergie est associé à un triplet dont toutes les composantes sont égales.

La transformation de l'espace RGB à l'espace XY Z peut s'interpréter comme un changement de base donné par la matrice :

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.166 & 0.125 & 0.093 \\ 0.060 & 0.327 & 0.005 \\ 0.000 & 0.004 & 0.46 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.2)$$

2.2.2.2. Les systèmes luminance-chrominance

La luminance est une composante commune à plusieurs systèmes de représentation de la couleur, ainsi que deux autres composantes de chrominances suffisantes pour quantifier le caractère chromatique d'un stimulus de couleur. Les systèmes de ce type peuvent être regroupés en une

CH II. Apport et utilisation des espaces de couleur dans la détection de la peau humaine sur des images.

famille particulière de systèmes de représentation de la couleur, les systèmes *luminance-chrominance*.

-Les systèmes luminance chrominance sont notés (L,Cr1,Cr2).

la notion de luminosité peut avoir plusieurs sens selon la façon dont elle est considérée et sera ainsi nommée, luminance, clarté, ou intensité.

Nous présentons ensuite les différents espaces de ce système :

2.2.2.2.1. Systèmes perceptuellement uniformes :

Les travaux de **MacAdam en 1942 [50]**, ont montré que dans le diagramme de chrominance une même distance entre deux couleurs ne sera pas perçue de la même façon par l'être humain selon la zone du diagramme considérée. Ainsi ,le système (X,Y,Z) n'est pas perceptuellement uniforme.

Ces systèmes possèdent une métrique permettant d'établir une correspondance entre une différence de couleur telle qu'elle est perçue par l'homme avec une distance dont la mesure reflète cette différence.

2.2.2.2.1.1. (U,V,W) :

proposé par la CIE sur la base des travaux de **Judd et de MacAdam [51]** en 1960 définissant un diagramme de chromaticité (u,v) CIE 1960 [CIE71]. Il est possible alors de définir le système (u,v,V).partir du système (U,V,W) de la même façon que le système(x,y,Y) est défini à partir du système (X,Y,Z). Cependant, ces premiers systèmes n'apportent pas de solutions Satisfaisantes au problème de mesures d'écart de couleurs et sont périmés.

2.2.2.2.1.2. LUV(1976) :

c'est une modification du système (U*V* W*) proposé par la CIE en 1964 , suite aux travaux de **Wyszecki [52]**, qui est d'ailleurs tridimensionnel en intégrant l'information de luminance, prenant ainsi en compte le type d'illuminant correspondant aux conditions d'éclairage. Il est noté (u*,v*,w*) , Le blanc de référence utilisé est alors caractérisé par ses composantes trichromatiques qui son Xw,Yw,Zw respectivement pour les primaires X,Y et Z.

2.2.2.2.1.3.CIELAB : CIELAB (L*a*b*)(1976):

Il a été développé par la CIE en 1976 dans le but d'améliorer la représentation des couleurs. Il s'agit du modèle chromatique le plus complet utilisé conventionnellement pour décrire l'ensemble des couleurs visibles pour l'oeil humain. Il s'agit d'un espace chromatique tridimensionnel dans lequel les différences de couleurs perçues comme étant d'importance égale sont également séparées par des distances égales. Cette différence peut être exprimée en delta-E

CH II. Apport et utilisation des espaces de couleur dans la détection de la peau humaine sur des images.

(DE). [53] Le delta-E est une description mathématique de la distance entre deux couleurs. Il fournit une mesure des variations de teinte et de densité.

Chaque couleur peut être précisément désignée à l'aide de ses valeurs "a" et "b" spécifiques et de sa luminosité - "L". Les trois paramètres du modèle représentent la luminance de la couleur - "L" (le niveau L le plus faible donne du noir), sa position entre le rouge et le vert - "a" (le niveau a le plus faible donne du vert) et sa position entre le jaune et le bleu - "b" (le niveau b le plus faible donne du bleu), en fonction d'un point blanc de référence.

Cependant, l'avantage de cet espace chromatique est son indépendance par rapport au matériel, et l'objectivité qui en résulte. La même combinaison de a, b et L décrit toujours exactement la même couleur.

Pour ces raisons, CIELAB est couramment utilisé comme référence au cours du processus de conversion des couleurs sur les systèmes ICC.

2.2.2.2.2. Les systèmes de télévision :

Les signaux émis par les chaînes de télévision sont reçus par les téléviseurs sous forme d'un signal unique appelé *signal composite*. Le récepteur décode ce signal composite sous forme de trois signaux primaires, appelés *primaires de synthèse*, par opposition aux *primaires d'analyse* ayant servi à coder le signal. À partir de ces signaux primaires, le récepteur effectue la synthèse additive de l'image couleur.

Il existe principalement deux types de standards :

NTSC pour les téléviseurs américains diffusant des images en 525 lignes.

PAL et SECAM pour les téléviseurs européens diffusant des images en 625 lignes.

2.2.2.2.2.1. Le système (Y',I',Q').

C'est le système défini par le standard NTSC.

ce standard n'utilise pas les primaires [Rc],[Gc],[Bc] , mais il est possible d'évaluer les composantes chromatiques Rc,Bc,Gc du standard NTSC à partir de celles de la CIE , en considérant l'illuminant C comme blanc de référence et en utilisant des matrices de passage.

2.2.2.2.2.2. Le système (Y',U',V') :

C'est le système défini par le standard PAL (norme allemande), Le blanc de référence utilisé par ce standard est l'illuminant D65. Il est aussi possible d'évaluer les composantes trichromatiques

CH II. Apport et utilisation des espaces de couleur dans la détection de la peau humaine sur des images.

Re,Ge,Be , d'un stimulus de couleur dans le système de primaires du standard PAL à partir du système (Rc,Gc,Bc) de la CIE. et réciproquement en utilisant une matrice de passage.

2.2.2.2.3. Autres systèmes :

Kodak a développé un système luminance-chrominance spécifique, appelé espace couleur photoYCC Kodak et noté (Y,C1,C2). La couleur des pixels des images acquises par un scanner photo CD Kodak est codée selon le système (Y,C1,C2) en vue du stockage sur des CD ROMs.

2.2.2.2.3. Les systèmes antagonistes :

Cette famille de systèmes de représentation de la couleur se base sur la théorie des couleurs opposées de Hering. Selon cette théorie, l'information couleur captée par l'oeil est transmise au cerveau sous la forme de trois composantes, une composante achromatique, A correspondant à la composante de luminance et deux composantes de chrominance, C1, C2.[54]

La composante A intègre les signaux fournis par les trois types de cônes de la rétine humaine et représente une opposition noir-blanc tandis que les composantes C1 et C2 intègrent uniquement les signaux fournis par certains types de cônes et correspondent respectivement à un signal d'opposition vert-rouge et à un signal d'opposition jaune-bleu. Notons que les systèmes (l^*,u^*,v^*) et (l^*,a^*,b^*) peuvent être considérés en ce sens comme des systèmes antagonistes. Afin de reproduire le fonctionnement du système visuel humain, quelques auteurs proposent différents modèles.

2.2.2.2.4. Autres systèmes :

2.2.2.2.4.1. Système de carron :

En étudiant les propriétés de différents systèmes luminance-chrominance, Carron passe du Système (R,G,B) défini par une caméra à un système luminance-chrominance noté (Y,Ch1,Ch2). Notons que la composante Ch1 du système de Carron correspond à une opposition cyan-rouge et la composante Ch2 à une opposition vert-bleu.

2.2.2.2.4.2. CIE xyY :

Afin de représenter efficacement une figure tridimensionnelle sur une feuille de papier bidimensionnelle, la CIE a transformé l'espace chromatique tridimensionnel en deux dimensions artificielles de couleurs (désignées collectivement par le terme chromaticité) et une d'intensité. Elle a ensuite prélevé une tranche bidimensionnelle dans cet espace, au niveau d'intensité maximal. Cette tranche est devenue le diagramme de chromaticité, également appelé "diagramme de chromaticité CIE xyY"

2.2.2.2.4.3. (Y,E,S) :

Système de luminance –chrominance déduit du standard (Rs,Gs,Bs) , du standard américain SMPTE (Society of Motion Picture and Television Engineers). Y correspond à l'information de luminance, E, représente une opposition vert-rouge et S représente une opposition bleu-jaune. Nous pouvons également citer l'utilisation d'un espace de couleur normalisé, noté (I,r,g), et composé d'une information de luminance I et des deux coordonnées trichromatiques r et g.

2.2.2.3. Les systèmes perceptuels :

l'homme ne perçoit pas la couleur comme une combinaison de composantes trichromatiques liées à des primaires mais selon des entités plus subjectives liées à la luminosité, la teinte et la saturation[55]. Il est donc naturel que de nombreux systèmes de représentation de la couleur quantifiant ces informations soient proposés dans la littérature. Nous les regroupons dans la famille des *systèmes perceptuels*.

Cette approche perceptuelle permet, en outre, une meilleure communication entre l'homme et la machine quand il s'agit d'exprimer la couleur. Il existe de nombreux systèmes de ce type dans la littérature, présentés sous différentes dénominations telles que ISH, HSL, HSV, TLS, LCH, LSH, LST, ITS, ... Ces différentes dénominations désignent les mêmes grandeurs mais se différencient par leur calcul. Nous distinguons deux familles de systèmes perceptuels : Beaucoup de systèmes perceptuels, plus compréhensibles par l'Homme, ont été créés. Certains d'entre eux se définissent en exprimant en coordonnées polaires ou cylindriques des composantes issues de systèmes luminance-chrominance. En l'occurrence, les espaces $L^*H^*C^*$ et $L^*H^*S^*$ sont respectivement issus d'une transformation des espaces $(L^*a^*b^*)$ et $(L^*u^*v^*)$. C, H et L désignent respectivement correspondent respectivement au chroma, à la teinte et à la luminance.

2.2.2.3.1. HSV Le modèle de cône hexagonal (pour Hue-Saturation-Value) :

La composante de teinte est matérialisée par un angle qui peut varier conventionnellement de 0 à 360°. Notons que c'est également le cas du système HLS. Les primaires additives et soustractives se situent sur les sommets des hexagones. La saturation S exprime l'éloignement de la couleur vis-à-vis de l'axe achromatique, tandis que la luminosité est donnée par la composante V, valant zéro pour le noir, et atteignant la valeur maximale au maximum de clarté que peut atteindre la couleur.

Principale avantage : assurer une bonne compatibilité avec la perception.

Inconvénients :

- l'uniformité assurée dans l'espace L^* , a^* , b^* (ainsi que l'espace L^* , u^* , v^*) ne l'est plus pour les espaces de luminance-chrominance.
- ils comportent certaines singularités. Notamment, la teinte n'est pas définie pour les valeurs non saturées.

2.2.2.4. Les systèmes d'axes indépendants :

Un des inconvénients des systèmes (R,G,B) est que ses trois composantes sont fortement corrélées. En effet, elles possèdent un fort facteur de luminance réparti sur chacune d'entre elles . Nous dirons que des composantes sont corrélées si elles portent en elle une information commune. Cette corrélation peut être plus ou moins importante selon les cas. Ainsi, de nombreux auteurs ont tenté de déterminer des systèmes de représentation de la couleur dont les composantes sont indépendantes, c'est-à-dire des composantes qui portent des informations différentes.

2.2.2.4.1. Le système d'Ohta :

Dans le domaine du traitement d'images couleur, une expérience menée par Ohta en 1980[56] sur une série de huit images différentes a permis de déterminer un système de représentation de la couleur basé sur la transformation de Karhunen-Loeve. Il montre alors qu'il existe un transformation unique appelée (I1,I2,I3).
modélisant cette transformation et se montrant tout aussi efficace que cette dernière.

$$I1 = \frac{R+G+B}{3}, \quad (2.3)$$

$$I2 = \frac{R-B}{3} \quad (2.4)$$

$$I1 = \frac{2*G-B-R}{3} \quad (2.5)$$

I1: plus discriminante représente, la luminance.

Les deux autres composantes représentent respectivement une opposition bleu-rouge et une opposition magenta-vert. Elles apportent moins d'information si bien que la troisième composante peut même être négligée car la valeur propre associée est très faible devant celle des deux autres.

2.3. Apports des espaces de couleur pour la détection de la peau humaine :

Le choix d'un espace de couleur dans un processus de détection de la peau peut être guidé par deux observations majeures sont :

- la couleur de la peau varie d'une personne à une autre, différentes études ont montré que la différence se situe plus au niveau de la composante de luminance qu'au niveau des composantes de chrominance.
- dans l'espace défini par les composantes de chrominance, les pixels correspondant à la peau sont "assez bien" regroupés.

Ces deux observations conduisent à privilégier les espaces de représentation qui séparent les composantes de luminance et de chrominance et à s'intéresser aux composantes de chrominance. L'étude comparative de 9 espaces de représentation par Terrillon [57] a montré que les meilleurs espaces de représentation des couleurs pour la détection de la peau sont les espaces normalisés *rgb* et *TSL*. L'espace *TSL* est un espace de représentation perceptuel qui sépare les composantes de teinte, de saturation et de luminance. Il est donc semblable à l'espace *HSI*. En nous basant sur cette étude, nous nous intéresserons à deux espaces de type luminance-chrominance, à savoir les espaces *rgb* et *YCrCb*, et à un espace perceptuel *HSI*.

Ce deuxième point est d'importance capitale puisque il informe de la capacité de ces espaces à séparer en représentation, la peau humaine de tout autre objet présent sur l'image .

La segmentation d'une image est possible grâce à une classification des pixels d'une image dans deux catégories : peau et non-peau.

Plusieurs approches sont possibles pour modéliser les zones de l'espace de chrominance caractérisant les pixels de la peau.

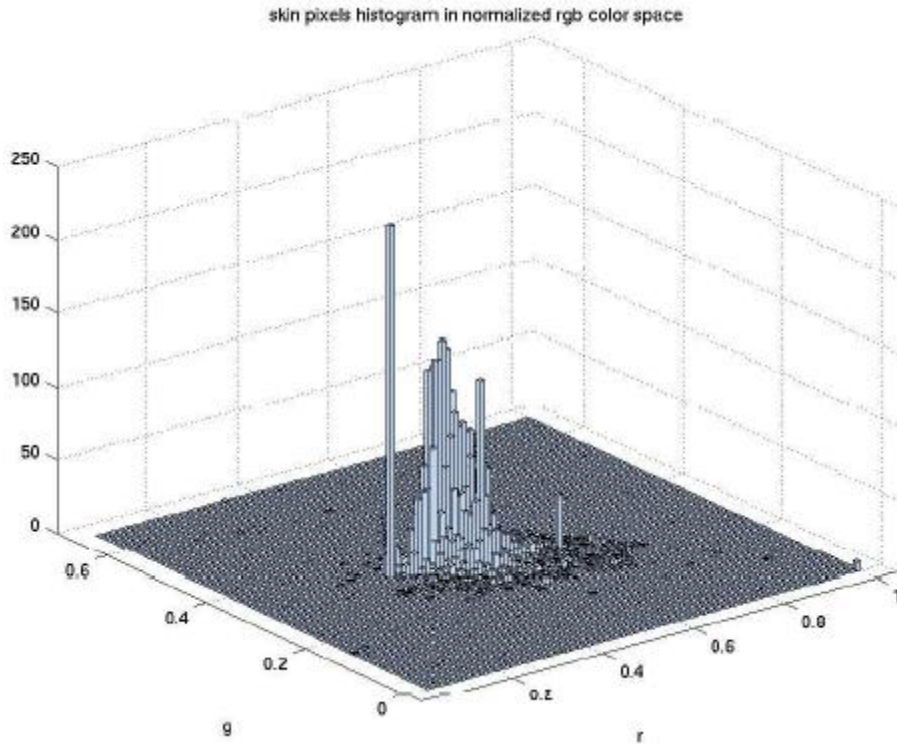


Fig.2.3. Histogramme des pixels de peau dans l'espace rgb . [49]

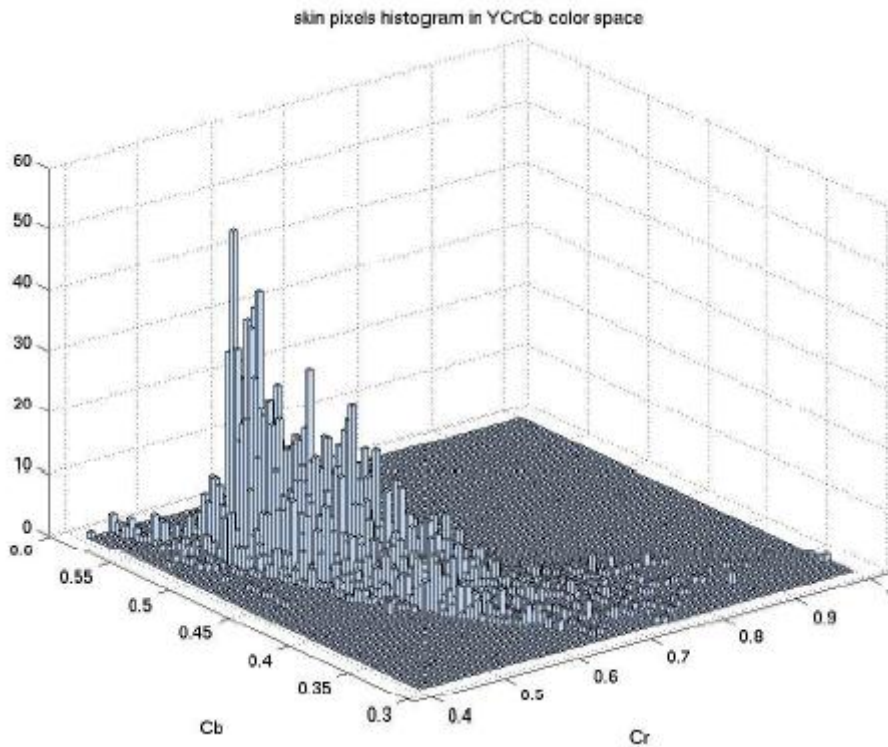


Fig.2.4. Histogramme des pixels de peau dans l'espace $YCrCb$. [49]

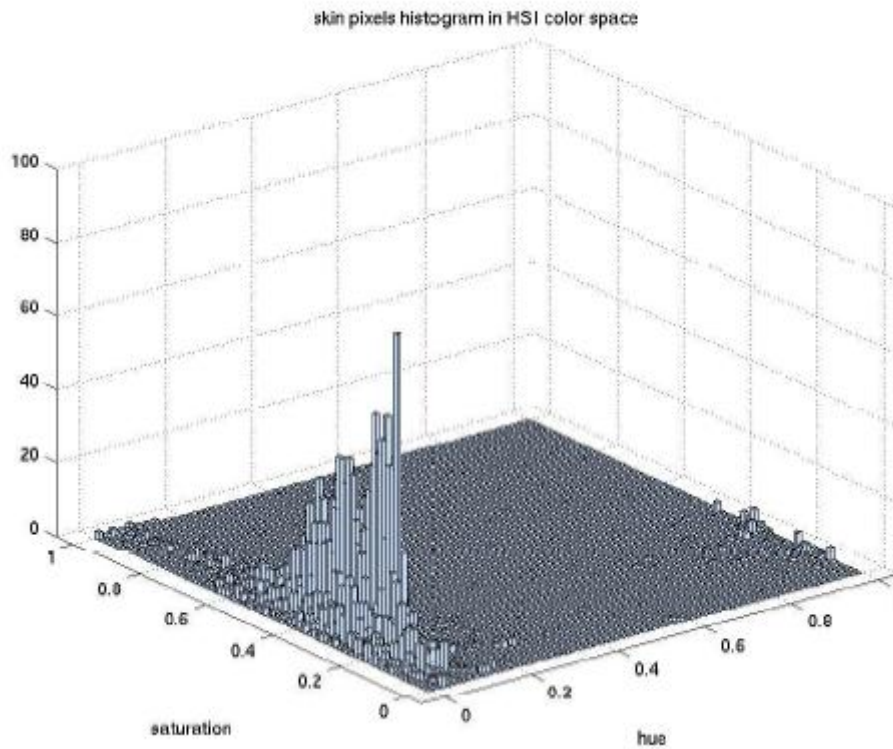


Fig.2.5. Histogramme des pixels de peau dans l'espace *HSI*. [49]

2.3. Conclusion :

Dans ce chapitre, nous avons présenté un ensemble de systèmes de représentation de couleurs les plus utilisés dans les industries.

Chacun ayant des propriétés particulières. Il faut savoir que cette liste est loin d'être exhaustive car il existe de nombreux autres systèmes parfois spécifiques à une application particulière.

Nous avons constaté que cette multitude de systèmes permet difficilement de les regrouper en différentes familles, ainsi qu'elle rend difficile le choix du système de représentation le mieux adapté à un problème donné.

Nous avons vu comment l'information colorimétrique présente dans une image couleur peut être utilisée pour la détection de la peau. Cela passe par le choix d'un espace de représentation convenable et par une modélisation adéquate de la distribution des couleurs dans cet espace,

CHAPITRE III.

Méthodes de détection de la peau humaine basée sur la couleur

3.1. Introduction :

L'exploitation de l'information de couleur est un moyen simple et efficace pour la discrimination entre les pixels de peau et ceux de non-peau. De nombreux chercheurs tels que, Fleck *et al.* [58] ,[59], Kjeldsen *et al.* [60], ont montré que la couleur de la peau est localisée dans une bande étroite de l'espace de couleur. Cette information peut donc facilement être utilisée pour marquer les pixels de couleur peau. .

Plusieurs modèles de couleur de peau ont été proposés pour la détection des pixels de peau humains dans les images couleur. La majorité de ces modèles utilisent une méthode de segmentation basée sur le calcul d'histogramme de couleur ou une méthode qui dérive de cette dernière [61].

On distingue principalement deux axes de recherche dans cette approche couleur. Le premier, concerne la modélisation de peau/non peau permettant de distinguer les pixels de couleur de peau de ceux de non-peau. Le deuxième axe concerne le choix de l'espace de couleur que nous devons adopter pour une bonne classification.

3.2. Fondements théoriques :

Les méthodes de classification se déclinent généralement en 2 familles :

On parle de mode supervisé quand on dispose d'un ensemble de points étiquetés pour l'apprentissage . Dans le cas contraire, nous parlons d'une classification non supervisée ou classification automatique.

Dans le cas où les classes sont connues (peau/non-peau), les méthodes de classification supervisée sont applicables. Ces dernières supposent la connaissance *a priori* de l'appartenance de chaque échantillon de l'ensemble d'apprentissage à une classe donnée, ce qui revient à supposer une connaissance *a priori* sur l'image à segmenter. En partant des données expérimentales que forment les échantillons tirés des classes, on obtient ainsi un effet d'apprentissage : à partir des échantillons soumis au système, ce dernier s'organise en vue de discriminer les échantillons ultérieurs. On dit que le système est capable de généraliser à partir des échantillons d'apprentissage.

On peut distinguer également deux catégories de méthodes de classification : les méthodes indirectes qui utilisent la formule de Bayes, et les méthodes directes qui évaluent les probabilités *a posteriori* sans faire intervenir la formule de Bayes. La formule de Bayes permet de déterminer les probabilités d'appartenance *a posteriori* si les densités de probabilité et les probabilités *a priori* sont connues. La règle de Bayes permet d'obtenir le taux d'erreur de classification minimum, ce qui est l'objectif souhaitable pour tout système de classification.

Les méthodes indirectes représentent la base de la majorité des travaux de recherche qui traitent la modélisation de la peau. A l'intérieur de ce groupe de méthodes, on distingue encore les méthodes paramétriques (qui font usage d'une hypothèse sur la forme analytique de la distribution) et les méthodes non paramétriques (qui ne font usage d'aucune hypothèse sur la forme de distribution).

Pour résumer, la conception d'un système de classification demande :

Pour résumer, la conception d'un système de classification demande :

1. La spécification des classes $\{Classe_k\}$.
2. La sélection des caractéristiques X des classes.
3. La spécification d'une représentation (modèle) pour $p(X/kClasse)$ et $p(Classe_k)$. Cette spécification peut être une fonction paramétrique ou non-paramétrique.
4. L'estimation des paramètres $P(X/Classe_k)$ à partir d'un ensemble $S_k = \{X_m\}$ de M_k observations.

Dans notre cas, pour chaque observation X , la classe $Classe k$ est connue, et peut être la classe peau ou la classe non-peau. Il suffit donc d'estimer les densités de probabilité.

3.2.1. Règle de décision de Bayes

Dans la mesure où les vecteurs de caractéristiques sont à N dimensions, les méthodes de la géométrie des espaces sont généralement utilisées. La classification se résume donc en une division de l'espace de caractéristiques en partitions disjointes. Cette division peut être faite par estimation de fonctions paramétriques ou par une liste exhaustive des frontières. Le critère généralement utilisé est la probabilité. Cette probabilité est fournie par la règle de Bayes. Nous rappelons l'expression mathématique de la formule de Bayes qui prend en considération la probabilité a priori d'apparition des individus des différentes classes et de leur distribution

$$P(Classe_k/X) = \frac{P(X/Classe_k)p(Classe_k)}{p(X)} \quad (3.1)$$

Avec : $P(Classe_k/X)$: probabilité a posteriori qu'un pixel de caractéristique X appartient à la classe k .

$P(X/Classe_k)$: Densité de probabilité de X si la classe est k .

$P(Classe_k)$: probabilité a priori qu'un pixel appartient à la classe k .

$P(X)$: Probabilité d'observer X .

La règle de décision de Bayes consiste à affecter l'individu à la classe dont la probabilité a posteriori (calculé par la formule de Bayes ou par toute autre méthode) est la plus grande. Cette décision minimise la probabilité d'erreur de classement. La figure 3.1 apporte une explication géométrique. Duda et Hart [62] donnent une preuve mathématique plus rigoureuse.

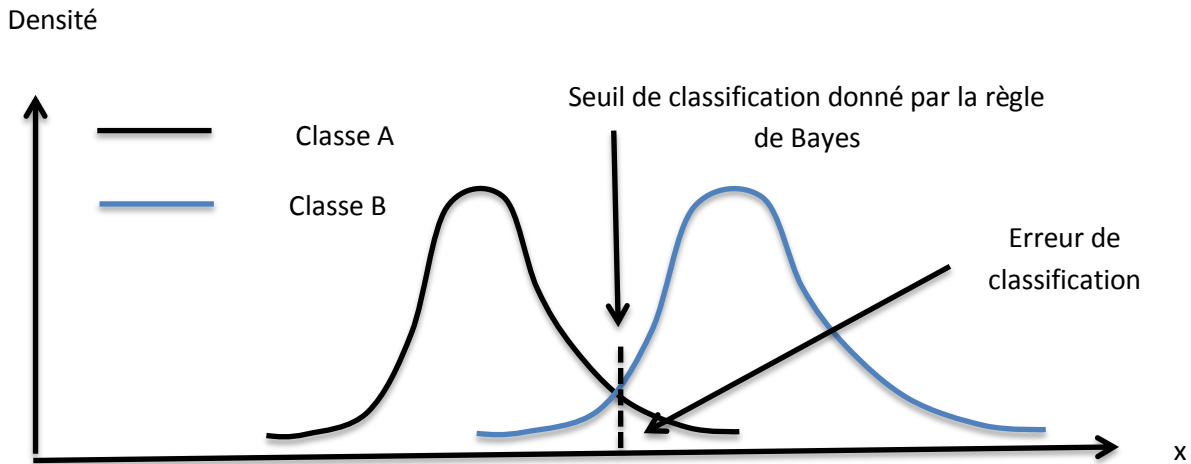


Figure 3.1. *Distribution des individus dans l'espace de description.*[61]

Les courbes représentent les fonctions densités de probabilité pondérées par les probabilités a priori correspondant aux deux classe A et B. De cette manière, ces courbes sont directement reliées à la densité des individus. Le trait vertical marque le seuil de classification donné par la règle de Bayes entre les deux classes.

L'erreur de classification est le nombre d'exemples A classée comme B et inversement. Elle correspond donc à la surface d'intersection des deux courbes. Si l'on choisit un autre seuil, nous nous apercevons que la nouvelle erreur de classification est égale à l'erreur de classification de Bayes augmentée d'une contribution positive. Elle est donc toujours supérieure à l'erreur de Bayes. Ainsi quel que soit le seuil pris pour séparer les 2 classes, l'erreur de classification est toujours supérieure à celle trouvée avec la règle de Bayes.

3.2.2. Estimation paramétrique des densités de probabilité

Les méthodes paramétriques consistent à faire une hypothèse concernant la forme analytique de la distribution de probabilité recherchée, et à estimer les paramètres de cette distribution à partir des données dont on dispose.

L'hypothèse la plus courante est que la répartition des individus de chacune des classes suit une loi gaussienne (figure 3.2). Elle conduit à la méthode appelée analyse discriminante avec une règle d'affectation probabiliste. Cette distribution « normale » des individus est la plus utilisée.

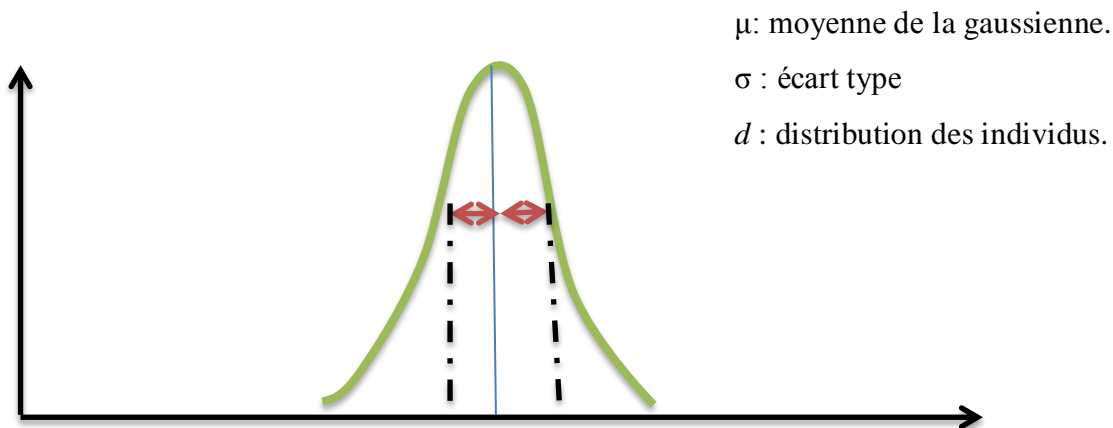


Figure 3.2. Distribution relative à une loi normale.[61]

3.2.2.1. Analyse discriminante avec une règle d'affectation probabiliste :

On rappelle l'hypothèse de distribution : les individus de la classe k sont répartis suivant une loi gaussienne multidimensionnelle :

$$P(x/Classe_k) = N(\vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{n/2} \sqrt{|\Sigma_k|}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \quad (3.2)$$

Avec Σ_k : matrice de covariance de la classe k .

μ_k : moyenne de la gaussienne de la classe k .

La matrice de covariance et le centre de la classe k sont estimés par la matrice de covariance et la moyenne des individus appartenant à la classe k .

Ainsi, à partir des estimations des matrices de covariance, des moyennes des gaussiennes (pour chacune des classes) et des probabilités *a priori*, on calcule par la formule de Bayes les probabilités *a posteriori* d'appartenance d'un individu aux classes. La règle de décision [61] consiste à classer un individu dans la classe qui obtient la plus grande probabilité *a posteriori*. La frontière de séparation est donc déterminée par l'ensemble des points pour lesquels les probabilités *a posteriori* sont égales.

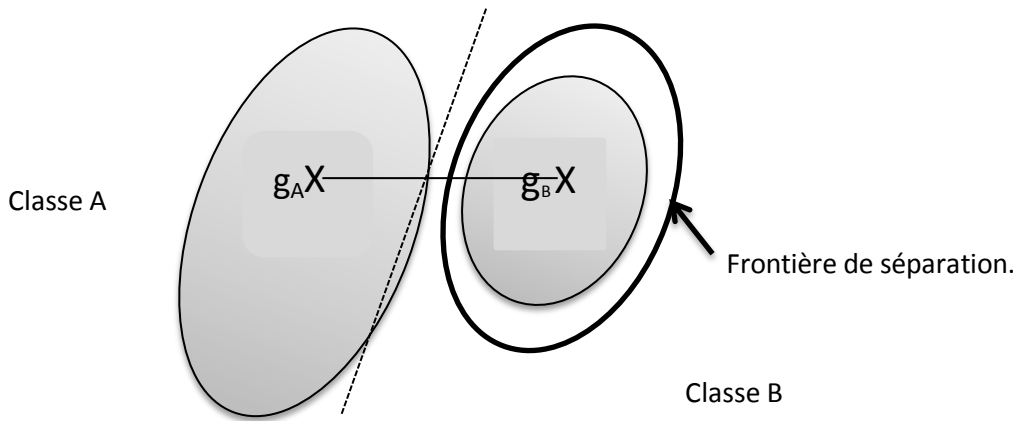


Figure 3.3. *Frontière de séparation*

(Analyse discriminante avec une règle d'affectation probabiliste). [61]

Sur la figure 3.3, on remarque que la frontière de séparation prend en considération la différence des classes. Le trait pointillé matérialise la frontière obtenue avec la règle d'affectation géométrique (voir section suivante).

Cette forme d'analyse discriminante peut sembler a priori très intéressante. Malheureusement même si les hypothèses sont vérifiées, les estimations des différentes matrices sont effectuées à partir des exemples qui risquent d'être peu nombreux. Elles sont donc très sensibles aux exemples marginaux.

3.2.2.2. Analyse discriminante avec une règle d'affectation géométrique [61] :

C'est la forme la plus simple de l'analyse discriminante. L'hypothèse de départ est complétée par les hypothèses suivantes pour garantir une convergence vers la règle de Bayes :

Les individus de la classe k sont répartis suivant une loi gaussienne multidimensionnelle déterminée par :

$$P(x/Classe_k) = N(\vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{n/2} \sqrt{|\Sigma_k|}} e^{-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k)} \quad (3.3)$$

avec Σ_k : matrice de covariance de la classe k ,

μ_k : moyenne de la gaussienne de la classe k .

- Les différentes matrices de covariance sont identiques :

$$\Sigma_1 = \Sigma_2 = \dots = \Sigma_c = \Sigma. \quad (3.4)$$

- Les probabilités a priori des classes sont, elles aussi, identiques :

$$Pr_1 = Pr_2 = \dots = Pr_c = 1/C \quad (3.5)$$

Dans ce cas, pour classer un nouvel exemple, l'analyse discriminante avec affectation géométrique calcule la distance (métrique de Mahalanobis) [44] entre l'exemple et les différents centres de gravités des classes, et affecte à cet exemple la classe correspondant à la plus petite distance. La distance de Mahalanobis (notée) est donc définie globalement dans l'espace de description par la matrice de covariance des individus :

$$\Delta_{\Sigma}^2 (\mu_1, \mu_2) = (\mu_1 - \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2) \quad (3.6)$$

La figure 3.4 montre un exemple de classification à deux classes :

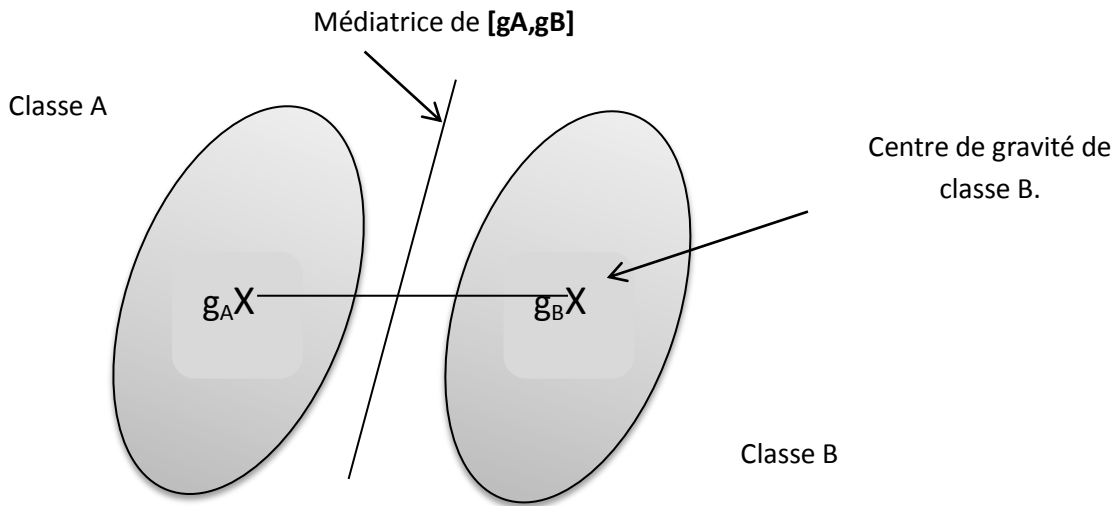


Figure 3.4. *Frontière de séparation*

(Analyse discriminante avec une règle d'affectation géométrique).[61]

Au sens de la métrique de Mahalanobis [63], la frontière entre les deux classes (A et B) est bien la médiatrice du segment $[g_A g_B]$.

Cette méthode de classification est très simple à mettre en oeuvre, car elle sépare les classes suivant des hyperplans (fonctions linéaires), malheureusement le résultat obtenu est rarement (voire jamais) celui que l'on obtiendrait par le classifieur de Bayes. Ainsi, une configuration typique est celle de la figure suivante :

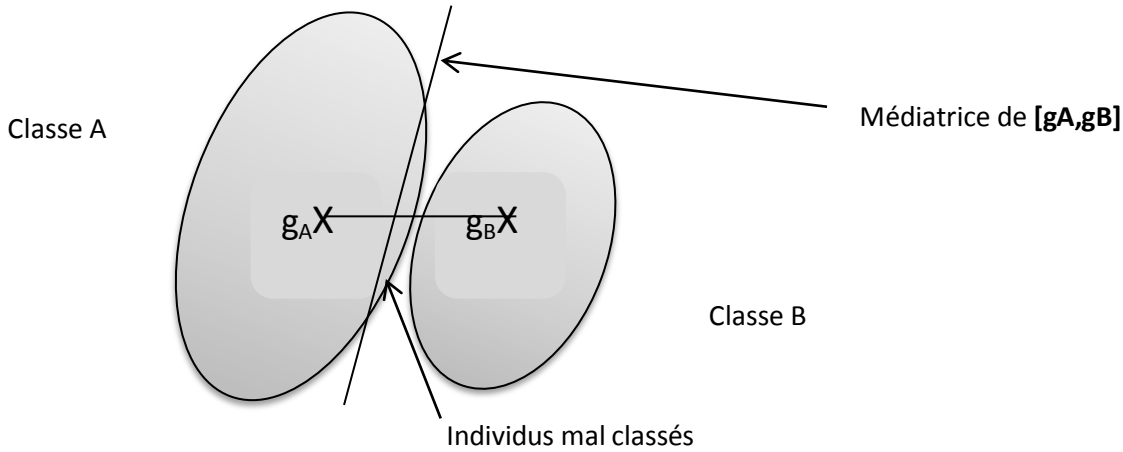


Figure 3.5. *Frontière de séparation*

(analyse discriminante avec une règle d'affectation géométrique).[61]

Ici, les individus de la classe A sont plus dispersés que ceux de la classe B. La frontière, quant à elle, n'a pas bougé par rapport à la figure 3.4 puisque les centres de gravité sont restés identiques. De nombreux individus de la classe A sont donc mal classés.

3.2.2.3 Estimation de la moyenne (μ) et la covariance (Σ) :

Nous rappelons que pour représenter $P(\vec{X}/Classe_k)$, les densités de probabilité de \vec{X} peuvent être estimées par une loi normale

$$P(\vec{X}/Classe_k) = N(\vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{n/2} \sqrt{|\Sigma_k|}} e^{-\frac{1}{2} (\vec{X} - \mu_k)^T \Sigma_k^{-1} (\vec{X} - \mu_k)} \quad (3.7)$$

L'apprentissage dans ce cas est réduit au problème d'estimation de la moyenne et de la covariance.

a. Estimation de la moyenne (μ) :

Soit M observations d'une variable aléatoire, $\{x_i\}$ La moyenne, μ_i , est l'espérance de $\{x_i\}$. Elle est

donnée par la formule:
$$\mu \equiv E\{x\} = \sum x p(x) \quad (3.8)$$

Pour les vecteurs de propriétés :

$$\vec{\mu} = E\{\vec{X}\} = \begin{pmatrix} \mu_1 \\ \dots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E\{x_1\} \\ \dots \\ E\{x_n\} \end{pmatrix} \quad (3.9)$$

Chaque composante est

$$\mu_i \equiv E\{x\} = \sum x_i p(x_i) \quad (3.10)$$

Cette moyenne peut être estimée par l'histogramme. On définit M la masse d'un histogramme $h(x)$ comme le nombre d'échantillons qui le composent.

$$M = \sum_{x_{min}}^{x_{max}} h(x) \quad (3.11)$$

On obtient ainsi :

$$\mu \equiv E\{x\} = \sum_{x_{min}}^{x_{max}} p(x) \cdot x \approx \frac{1}{M} \sum_{x_{min}}^{x_{max}} h(x) \cdot x \quad (3.12)$$

Pour un vecteur de variables à n dimensions, on aura :

$$\vec{\mu} \equiv E\{X_n\} = \sum_{x_{min}}^{x_{max}} p(\vec{X})x_n \approx \frac{1}{M} \sum_{x_{min}}^{x_{max}} h(\vec{X})x_n \quad (3.13)$$

b. Le deuxième moment : la covariance

La variance est le deuxième moment de la densité de probabilité. σ^2

Pour un ensemble d'observations M , la variance de la variable x_i est :

$$\sigma^2 \equiv E\{(x - \mu)^2\} = \frac{1}{M} \sum_{i=1}^M (x_i - \mu)^2 \quad (3.14)$$

Mais l'usage de μ estimé avec le même ensemble, introduit un biais dans !!. Pour l'éviter, on utilise une estimation sans biais.

$$\sigma^2 = \frac{1}{M-1} \sum_{i=1}^M (x_i - \mu)^2 \quad (3.15)$$

Avec μ et σ^2 on peut estimer la densité $p(x)$ par :

$$P(x) = N(\vec{\mu}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \quad (3.16)$$

σ^2 est le deuxième moment de $p(x)$.

$$\sigma^2 \equiv E\{(x - \mu)^2\} = \sum_{x=x_{min}}^{x_{max}} P(x) \cdot (x - \mu)^2 = \frac{1}{M} \sum_{x=x_{min}}^{x_{max}} h(x)(x - \mu)^2 \quad (3.17)$$

Pour N dimensions, la covariance entre les variables x_i et x_j est estimée à partir de M observations

$$\sigma_{ij}^2 \equiv E\{(x_i - E\{x_i\})(x_j - E\{x_j\})\} \quad (3.18)$$

$$= \frac{1}{M} \sum_{m=1}^M (x_{im} - \mu_i)(x_{jm} - \mu_j) \quad (3.19)$$

et encore pour éviter le biais, on utilise :

$$= \frac{1}{M-1} \sum_{m=1}^M (x_{im} - \mu_i)(x_{jm} - \mu_j) \quad (3.20)$$

Ces coefficients composent la matrice de covariance Σ :

$$\Sigma_x \equiv E\{(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})^T\} = E\{(\vec{X} - E\{\vec{X}\})(\vec{X} - E\{\vec{X}\})^T\} \quad (3.21)$$

$$\Sigma_x \equiv \begin{pmatrix} \sigma_{11}^2 & \sigma_{11}^2 & \dots & \sigma_{11}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2n}^2 \\ \dots & \dots & \dots & \dots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \dots & \sigma_{nn}^2 \end{pmatrix} \quad (3.22)$$

On obtient les paramètres de la fonction :

$$P(\vec{X} / \text{Classe}_k) = N(\vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{n/2} \sqrt{|\Sigma_k|}} e^{-\frac{1}{2}(\vec{X}-\vec{\mu})^T \Sigma_k^{-1} (\vec{X}-\vec{\mu})} \quad (3.23)$$

3.2.3. Estimation non paramétrique des densités de probabilité

Lorsqu'on ne peut pas faire d'hypothèse sur la distribution des individus, il faut se tourner vers des méthodes non paramétriques. Il s'agit de s'affranchir de l'hypothèse d'une loi paramétrique et d'estimer la distribution $P(x) = P(X=x | K=i)$ de la classe courante par approximations.

Le principe de l'estimation non paramétrique de la densité de probabilité est de délimiter une région R_N autour d'un point considéré, puis de compter le nombre d'individus dans ce volume, et enfin de déterminer la densité comme le rapport entre ce nombre (divisé par le nombre total d'individus) et le volume de la région ([64], [65] et [66]). Ainsi, on obtient une estimation de la densité de probabilité avec la formule suivante :

$$\hat{P}_N(x) = \frac{k_N}{N.V_N} \quad (3.24)$$

Avec :

N : nombre d'individus de l'échantillon

k_N : nombre d'individus dans la région R_N

V_N : volume de R_N

Ayant une telle fonction, on peut utiliser une technique optimale de classification en appliquant la règle de Bayes.

Le problème principal dans l'estimation d'une fonction de densité pour un vecteur de caractéristiques est la croissance exponentielle du nombre de cellules N avec le nombre de dimensions D . Ceci induit une croissance exponentielle dans les nombres d'exemples M nécessaires. En Anglais, on appelle ce problème "The Curse of Dimensionality".

3.3. Modèle de peau paramétrique

Les modèles de peau paramétriques permettent d'ajuster les distributions avec quelques fonctions spécifiques paramétrées. Ils offrent trois avantages : (1) ils apportent un gain en espace mémoire ainsi qu'en possibilité de manipulation ; (2) ils donnent plus d'intelligence et de finesse dans les vraies formes ou dans la régularité des distributions ; et (3) ils ont la capacité

d'interpoler les données d'apprentissage quand elles sont dispersées. Différentes fonctions peuvent être appliquées en fonction du problème.

3.3.1. Modèle basé sur une simple gaussienne :

Cette méthode a été utilisée par [67], [68] et [69]. La distribution de couleur de peau est estimée par une fonction de densité de probabilité gaussienne :

$$P(c|peau) = \frac{1}{2\pi \sqrt{|\Sigma_{peau}|}} e^{-\frac{1}{2}(c-\mu_{peau})^T \Sigma_{peau}^{-1} (c-\mu_{peau})} \quad (3.25)$$

où c est la variable aléatoire à deux dimensions représentant le couple de chrominance; μ_{peau} et Σ_{peau} sont respectivement l'espérance et la matrice de covariance représentant les paramètres du modèle gaussien. Ces paramètres peuvent être estimés à partir de l'échantillon d'apprentissage selon les équations suivantes :

$$\mu_{peau} = \frac{1}{N_{peau}} \sum_{c \in C} N_{peau}(c)c \quad (3.26)$$

$$\Sigma_{peau} = \frac{1}{N_{peau}-1} \sum_{c \in C} N_{peau}(c)(c-\mu_{peau})(c-\mu_{peau})^T \quad (3.27)$$

3.3.2. Modèle basé sur un mélange de Gaussiennes.

D'autres auteurs [70], [71], [72], [73], [74] et [75] ont également proposé l'utilisation d'un modèle mixte de gaussiennes (Gaussian Mixture Model), qui vise à mieux représenter et modéliser la portion d'un espace de couleur associée à la couleur peau .

En général les densités de probabilités sont plus complexes que $N(\mu, \Sigma)$. Une fonction paramétrique plus générale est un mélange de Gaussiennes.

Le mélange des Gaussiennes est une extension des gaussiennes simples. A la différence d'une gaussienne simple, il a la capacité de représenter les distributions les plus complexes. Dans ce cas la fonction de densité de probabilité gaussienne est représentée comme suit:

$$P(c|peau) = \sum_{n=1}^N w_n p_n(c|peau) \quad (3.28)$$

Avec :

p_n : les noyaux des gaussiens définis dans (3.27). Chaque p_n est lui-même une distribution gaussienne.

N est le nombre de noyaux gaussiens qu'il faut choisir correctement de sorte que le modèle puisse bien représenter les données d'apprentissage

w_n sont les poids des noyaux correspondants dont la somme est égale à 1.

Dans ce cas, trois paramètres sont à estimer (w_n, μ_n, Σ_n). Généralement l'algorithme EM (Expectation Maximization) est appliqué pour estimer ces paramètres.

La figure 3.6 illustre un mélange de trois gaussiennes en proportions égales (à gauche) et l'histogramme correspondant (à droite). La courbe en trait plein à gauche représente la densité de probabilité théorique du mélange résultant en tenant compte des proportions.

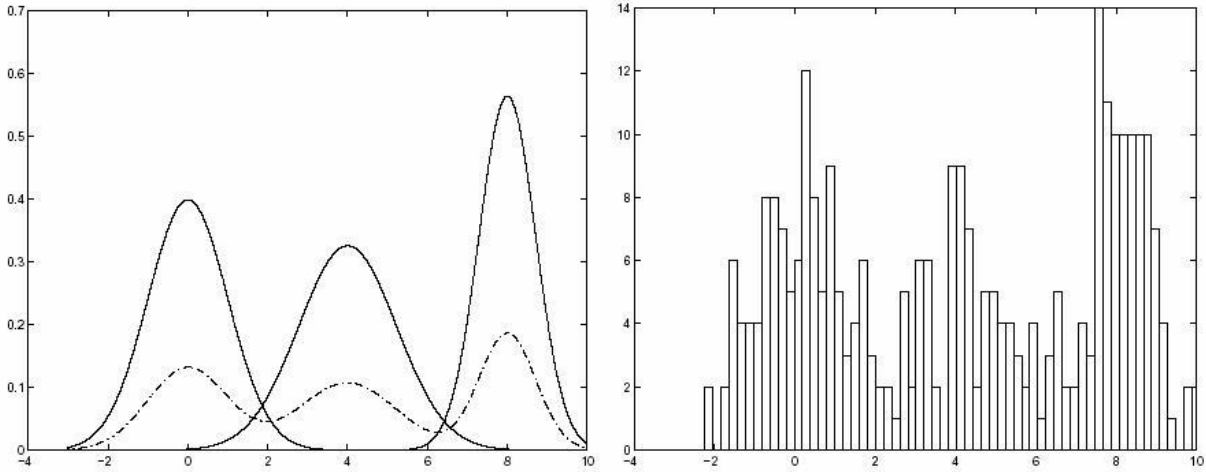


Figure 3.6. (A gauche) un mélange de 3 gaussiennes (en pointillé) et leurs densités théoriques (trait plein) (A droite) l'histogramme associé. [61].

3.3.3 Modèle elliptique de borne :

En examinant la distribution de couleur de peau et de non-peau dans plusieurs espaces de couleur, Lie et Yoo [76] ont conclut que cette distribution prend approximativement la forme d'une ellipse. Ils ont séparé les régions de couleur de peau et de non-peau par une borne elliptique afin de caractériser la vraie forme de la distribution des données d'apprentissage. Le modèle est défini comme suit:

$$\Phi(c) = (c - \varnothing)^T \Sigma_{peau}^{-1} (c - \varnothing) \quad (3.29)$$

Où \varnothing et Σ_{peau} sont calculés à partir des données d'apprentissage comme suit :

$$\varnothing = \frac{1}{|C_{peau}|} \sum_{c \in C_{peau}} c \quad (3.30)$$

$$\Sigma_{peau} = \frac{1}{N_{peau}} \sum_{c \in C_{peau}} N_{peau}(c) (c - \mu_{peau})(c - \mu_{peau})^T \quad (3.31)$$

où $|C_{peau}|$ est le nombre de pixels de couleur de peau avec C_{peau} . μ_{peau} est l'espérance des vecteurs d'apprentissage de pixels de peau défini par l'équation (3.26).

Pour classifier les pixels on compare $\Phi(c)$ avec un seuil fixé θ . c est considéré comme un pixel de peau si $\Phi(c) < \theta$ sinon il est un pixel de non peau.

Les taux de classification obtenus par ce modèle sur six espaces de couleur testés sont satisfaisants. De plus, cette méthode est plus rapide lors de la classification que les méthodes basées sur une simple gaussienne [77].

3.4. Modèle de peau non paramétrique

La recherche selon cette approche vise à estimer la distribution de couleurs de peau des données d'apprentissage sans aucune hypothèse sur la distribution.

3.4.1 Classification de pixels par table de correspondance : lookup table (LUT)

Cette méthode consiste à construire dans un premier temps un modèle de peau en se basant sur le calcul d'histogramme à partir d'exemples. Elle l'applique ensuite à une image ou à une scène pour obtenir une carte de probabilité en associant à chaque pixel sa probabilité d'appartenance à une instance du modèle. A partir de cette carte, les pixels sont classés. Cette méthode a été appliquée dans plusieurs algorithmes de détection et de suivi de visage pour segmenter les pixels de peau ([78], [79], [80],[81], [82], [83] et [84]).

Après projection dans un espace de couleur bien choisi, le modèle prend la forme d'un histogramme des valeurs des pixels de la base d'exemples. En effet, l'espace de couleur est quantifié par un nombre de case (bins). A chaque case des axes utilisés est associé le nombre de fois que la valeur de couleur s'est produite dans les images de peau de la base d'apprentissage. Ces cases forment donc un histogramme de couleur (3D) ou (2D) désigné dans certains travaux sous le nom de table de correspondances « Lookup Table ».

Nous signalons que les histogrammes (2D) ont été plus utilisés dans la littérature. Nous citons l'espace (r, g) (composantes R et G normalisées) par Bérard [85], l'espace (H, S) (teinte et saturation) par Wu [66]. En effet, à partir d'un histogramme tri-dimensionnel (3-D) on peut déduire trois histogrammes (2-D), par projection de l'histogramme (3-D) sur 2 des 3 plans colorimétriques d'une image I . Un histogramme (2-D) peut donc être considéré comme une image J dont les dimensions spatiales sont celles des deux axes choisis dans l'espace colorimétrique associé à I .

D'autres chercheurs préfèrent l'utilisation d'un histogramme tridimensionnel. Quek *et al.* [86] ont construit leur table de correspondances dans l'espace RGB à partir d'une base d'images prise contre un arrière plan simple afin de permettre une segmentation facile des pixels.

Une fois l'histogramme construit, il est normalisé et ces valeurs sont converties en distribution discrète $p(c|peau)$

$$p(c|peau) = \frac{peau(c)}{Norm} \quad (3.32)$$

où $peau(c)$ est le nombre de pixels associé à une case (bins) de l'histogramme de peau formé par le vecteur de couleur c et $Norm$ est le coefficient de normalisation.

Ce coefficient ($Norm$) a été fixé selon différentes manières. Il correspond au nombre total de pixels de peau dans les travaux de [85] et [86]. Il est égal à la plus grande valeur dans la table de correspondance (Lookup Table) dans les travaux de [77]. Les valeurs normalisées de cette table constituent la probabilité qu'une couleur correspond à la peau. Un pixel est considéré comme pixel de peau si sa valeur dépasse un certain seuil.

3.4.2 Modèle basé sur l'appariement d'histogrammes

Saxe et Foulds [87] utilisent l'espace de couleur HSV et une technique basée sur l'appariement d'histogrammes, décrite dans [88]. L'histogramme de couleur calcule la fréquence d'apparition de chaque couleur. Cette information est globale et elle ne prend pas en compte la disposition spatiale des pixels dans l'image. Elle est donc invariable à la rotation et à la translation, ce qui est une propriété intéressante pour apparier. Les auteurs suggèrent une sélection d'une région de peau, appelée région de contrôle, qui est comparée ensuite au reste de l'image en utilisant l'appariement d'histogramme. Il s'agit de comparer l'histogramme de la région de contrôle avec les histogrammes des régions de même taille de l'image. Dans un premier temps l'algorithme transforme l'image d'entrée dans l'espace de couleur HSV. Puis c'est à l'utilisateur de choisir manuellement le germe initial (région de contrôle). Une fois le bloc de contrôle choisi, l'image est examinée région par région en utilisant la méthode d'intersection d'histogramme (équation 3.33).

$$M_{C,I} = \frac{\sum_{i,j} \min(H^C(i,j), H^I(i,j))}{\sum_{i,j} H^C(i,j)} \quad (3.33)$$

3.4.3 Modèle Bayésien basé sur les Histogrammes

Dans [89] [90], les auteurs présentent leur modèle de couleur de peau et de non-peau par des histogrammes. Ils quantifient l'espace de couleur C à un certain nombre de bins $c \in C$ et comptent le nombre de pixel de couleur dans chaque bins. $N_{peau}(c)$ représente donc le nombre de pixels de couleur c pour la classe de peau et $N_{-peau}(c)$ pour la classe de non-peau. Enfin ils normalisent chaque bin pour obtenir la distribution conditionnelle des pixels de peau et de non peau $p(c/peau)$ et $p(c/non\ peau)$. Supposons que N_{peau} représente le nombre total de pixel de peau et N_{-peau} le nombre total de pixels de non-peau dans la base d'apprentissage, nous avons :

$$p(c/peau) = \frac{N_{peau}(c)}{N_{peau}} \quad (3.34)$$

$$p(c|\neg\text{peau}) = \frac{N_{\neg\text{peau}}(c)}{N_{\neg\text{peau}}} \quad (3.35)$$

Ce qui signifie aussi :

$$p(\text{peau}) = \frac{N_{\text{peau}}}{N_{\text{peau}} + N_{\neg\text{peau}}} \quad (3.36)$$

$$p(\neg\text{peau}) = \frac{N_{\neg\text{peau}}}{N_{\text{peau}} + N_{\neg\text{peau}}} = 1 - p(\text{peau}) \quad (3.37)$$

Par la suite, la formule de Bayés ci dessous est utilisée pour calculer la probabilité qu'un pixel soit un pixel de peau ou non selon sa couleur.

$$P(\text{peau}/c) = \frac{p(c|\text{peau})p(\text{peau})}{p(c|\text{peau})p(\text{peau}) + p(c|\neg\text{peau})p(\neg\text{peau})} \quad (3.38)$$

$$P(\neg\text{peau}/c) = 1 - p(\text{peau}/c) \quad (3.39)$$

La décision concernant les deux classes est faite selon un seuil choisi Θ , $0 < \Theta < 1$. Le pixel sera considéré comme un pixel de peau si $p(\text{peau}/c) > \Theta$ et comme un pixel de non-peau si $p(\text{peau}/c) \leq \Theta$.

3.4.4 Self-organizing map (SOM) :

Décrit par Kohonen au début des années 80, le réseau SOM (Self-Organizing Map) [91] est un des plus populaires réseaux de neurones non supervisé. Il est principalement utilisé pour classifier des données de grande dimension, mais il fonctionne tout aussi bien pour des données de faible dimension. L'objectif du réseau SOM est de classifier les pixels d'entrée selon qu'ils correspondent ou non à des pixels de peau, ce qui permet au final d'extraire la région du visage. Brown et al. [92] entraînent deux réseaux SOM pour apprendre, à partir de 500 images, la distribution des pixels de la couleur peau et la couleur non peau. Les performances du SOM ont été testées sur l'ensemble des images d'apprentissage/test de la base de données Compaq [93]. Plusieurs espaces couleur (RGB normalisé, Teinte-Saturation, TLS) ont été utilisés avec le détecteur SOM. Les résultats ont montré que les performances des détecteurs de peau SOM ne dépendent pas des espaces couleurs utilisés, à l'inverse du modèle MMG. Par ailleurs, les performances du détecteur SOM sont inférieures à celles qui sont obtenues par la méthode basée sur les histogrammes RGB, développée dans [93]. Néanmoins, le détecteur SOM nécessite moins de paramètres et il est plus efficace pour les applications en temps réel. Certains auteurs ont également proposé un apprentissage automatique des couleurs représentant la peau à l'aide de réseaux neuronaux. Ceux-ci peuvent être entraînés à partir d'échantillons de pixels représentant la peau (et non-peau), préalablement convertis dans

l'espace de couleur YCrCb. Par ailleurs, d'autres espaces de couleurs ont aussi été utilisés dans ce contexte. Kakumanu et al. [94][95] ont employé un réseau de neurones pour modéliser la constance de la couleur « the color constancy ». Leur réseau de neurones possède trois couches et permet d'estimer l'illumination de la peau. Il prend en entrée un histogramme à deux composants r et g et donne en sortie les illuminations de la peau dans l'espace r, g. Ce réseau de neurones est entraîné sur une base de données de 255 images, et testé sur 71. Les images utilisées fournissent un grand choix d'illuminations, de différents arrière plans et de sources de lumière non blanches. Enfin, une technique de seuillage simple est utilisée pour détecter la peau à partir de ce réseau de neurones.

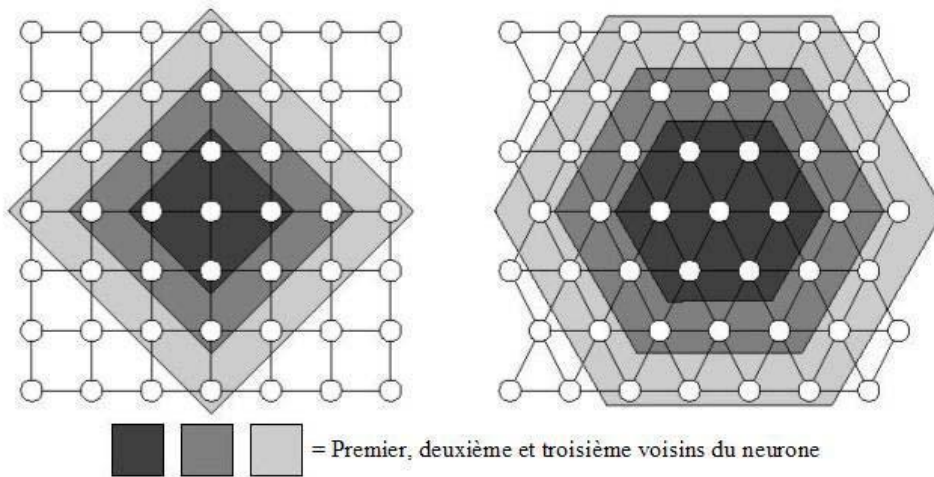


Figure 3.7. Architecture possibles de la carte SOM : rectangulaire (à gauche) et hexagonale (à droite).[61]

3.5 Autres modèles basée sur la couleur

Dans des applications spécifiques telles que le suivi de visage, ou encore le filtrage des images de nudités, la détection de peau est juste une étape préalable et devrait donc être rapide. Une méthode pour établir un classifieur de peau rapide est de définir explicitement (par un certain nombre de règles) les bornes des régions de peau dans un espace de couleur. A titre d'exemple Peer *et al.* [96] considèrent un pixel comme pixel de couleur de peau si chacune des conditions suivantes est respectée :

$$\left\{ \begin{array}{l} R > 95, G > 40, B > 20 \\ \text{Max}\{R, G, B\} - \text{min}\{R, G, B\} > 15 \\ |R - G| > 15, R > G, R > B \end{array} \right. \quad (3.40)$$

L'avantage est la simplicité des règles de décision utilisées pour discriminer les pixels de peau de ceux de non peau, ce qui produit une classification très rapide. Néanmoins, l'inconvénient majeur consiste dans le choix de l'espace de couleur ainsi que les règles et les seuils de décision par des études empiriques.

3.6 Espaces de couleur utilisés pour la modélisation de la peau

Dans les systèmes de vision, le choix d'un système adapté de la représentation de la couleur est un problème très délicat [97]. L'espace de couleur le plus fréquemment utilisé est celui de RGB. Il décrit la couleur comme la corrélation de trois couleurs primaires (Rouge-Verte-Bleue). Chaque pixel d'une image couleur contient donc les trois composantes qui définissent sa couleur, C'est le principe de la vision trichromatique.

D'autres espaces de représentation de couleurs peuvent être dérivées par des combinaisons linéaire ou non linéaire des composantes R, G et B. Ces espaces de couleur ont été proposés dans différentes applications [98] dont plusieurs ont été appliqués au problème de la modélisation de la couleur de la peau.

Des tests comparatifs, sur le meilleur espace de couleur permettant de détecter les zones de peau dans une image, ont été effectuées par Nicolas MOTTIN [99]. Cinq espaces de couleur différents ont été testés : YUV, RGB, HSI, TSL, et Lab. Les tests ont été faits sur la base de peau de Von Luschan (cf. figure 3.10), qui contient 36 imagettes de peau depuis les peaux les plus claires jusqu'aux peaux les plus mates. Ces résultats ont montré que la meilleure détection de peau peut se faire dans l'espace HSI.

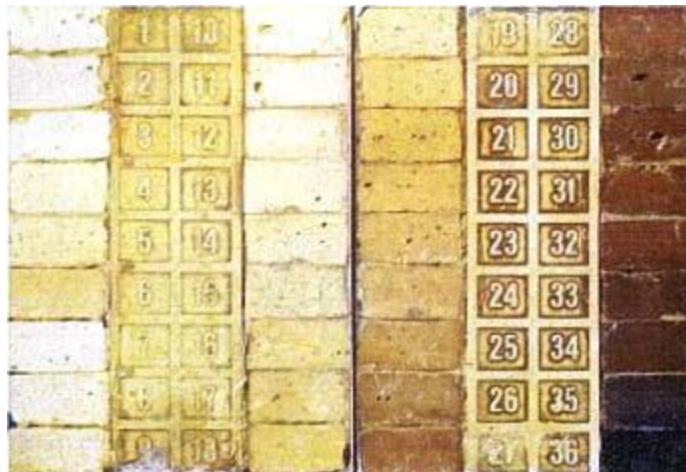


Figure 3.8. Base de peaux de Von Luschan.[61].

3.7. Conclusion :

Dans ce chapitre, nous avons fourni la description, la comparaison et les résultats de l'évaluation des méthodes populaires pour la modélisation et la détection de la peau humaine.

Nous avons essayé de résumer les importantes différences entre les méthodes, ainsi que leurs avantages et leurs inconvénients.

Les conclusions les plus importantes que nous en tirons sont énumérées ci-dessous:

Dans les méthodes basées sur la couleur, on distingue encore les méthodes paramétriques et les méthodes non paramétriques, ainsi que des méthodes basées sur une fixation empirique des règles et des seuils de décision. Pour les modèles paramétriques, leur pertinence et leur de l'espace de couleur choisi. Par ailleurs, les phases d'apprentissage et de test pour ce type de méthodes sont lentes puisque celles-ci impliquent une procédure d'évaluation des paramètres tel que l'algorithme (EM) et l'évaluation des fonctions relativement complexes comme (2.29), (2.32) et (2.33).

Quant aux méthodes basées sur la fixation empirique des règles et des seuils de décision, elles sont des méthodes simples et rapides ce qui explique leurs utilisations pour des applications qui fonctionnent en temps réel. Néanmoins, on peut reprocher à ces méthodes la façon dont les règles de décision et les seuils sont choisis.

Les méthodes non paramétriques sont généralement rapides dans la phase d'apprentissage et de test. Elles ne font aucune hypothèse sur la répartition des données d'apprentissage et elles sont théoriquement indépendantes de la forme de distribution de la peau, contrairement aux méthodes paramétriques.

CHAPITRE IV.

L'astuce noyau (kernel trick).

4.1. Introduction :

Aujourd'hui, les systèmes conçus pour vérifier si un billet est vrai ou faux, pour reconnaître une personne à partir du son de sa voix ou de son visage ou pour détecter les données aberrantes dans une base de données intègrent des algorithmes issus de la théorie statistique de l'apprentissage et en particulier une machine à noyaux comme outil de décision. La programmation de ces systèmes utilise un ensemble de **couples observation-étiquette** pour élaborer une règle de décision. On parle alors d'apprentissage statistique ou de **programmation par l'exemple**[100]. Les machines à noyaux constituent une classe d'algorithmes permettant d'**extraire de l'information à partir de données dans un cadre non paramétrique**. L'intérêt suscité par ces méthodes tient d'abord aux excellentes performances qu'elles ont permis d'obtenir notamment sur les problèmes de grande taille. Cette bonne tenue à la charge est due à la parcimonie de la solution et à la faible complexité de son calcul. L'intérêt des machines à noyaux réside aussi dans leur caractère flexible et rigoureux, approche, qui recèle un grand potentiel. Dans ce qui suit nous allons présenter les machines à noyaux en se focalisant sur la , la régression logistique à noyau ,en faisant le point sur les différentes facettes de son utilisation. L'accent est mis sur les considérations pratiques liées à la mise en œuvre de ce type de méthode.

4.2. Noyaux :

4.2.1. L'espace des caractéristiques (feature space) [101] :

Imaginons un ensemble d'exemples x_j étiquetés par -1 ou 1 suivant la classe à laquelle ils appartiennent, qui ne soit pas du tout séparable linéairement. La séparation linéaire peut fonctionner , mais la séparation est bien entendu de piètre qualité, et bon nombre de vecteurs sont des supports (*Fig .4.1*).

Une solution pour mieux séparer les exemples est de les projeter dans un espace différent , et de réaliser une séparation linéaire dans cet espace-là, où cette fois-ci elle devrait être plus adaptée [102].

Soit ϕ cette projection, on a :

$$\phi(\vec{x}) = \begin{pmatrix} \phi_1(\vec{x}_1) \\ \phi_2(\vec{x}_2) \\ \vdots \\ \phi_n(\vec{x}_n) \end{pmatrix} \quad (4.1)$$

bien sûr, les fonctions ϕ_i sont pas nécessairement linéaires, et on peut même avoir $n = \infty!$.

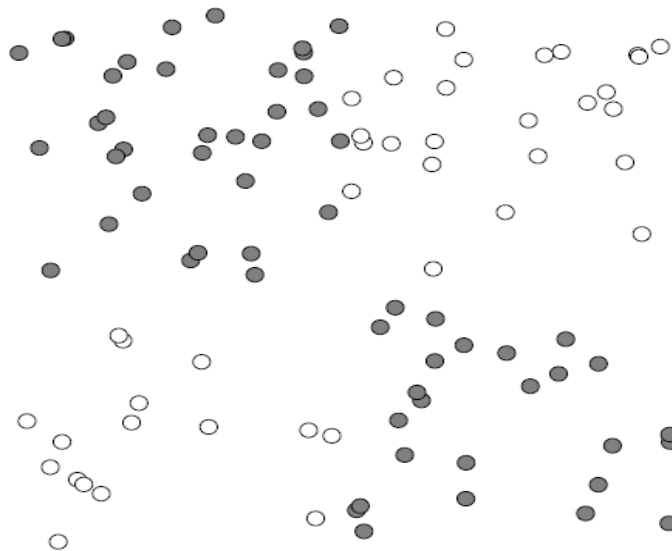


Fig.4.1. Cas non linéairement séparable.[102]

Le gros intérêt des noyaux est que tout ce qui s'applique sur la séparation linéaire s'applique en fait très facilement à des séparations non linéaires, sous réserve de bien faire les choses. [103] [104]

4.2.2. Définition : (Fonction noyau (Kernel function)) [105]

Une fonction noyau est une fonction $k : x, x' \in \mathbb{R}^n$ satisfaisant :

$$k(x, x') = (\phi(x), \phi(x')) \quad (4.2)$$

où ϕ est une fonction de X vers un espace de redescription F doté d'un produit scalaire :

$$\phi \rightarrow (\phi) \in F \quad (4.3)$$

L'utilisation de fonctions noyau permet ainsi de calculer implicitement un produit scalaire dans un espace de dimension éventuellement infini par un calcul n'impliquant qu'un nombre fini de termes, ce nombre étant le nombre m des exemples d'apprentissage (voire éventuellement moins, comme nous le verrons). Par ailleurs, il apparaît donc que la spécification de la fonction noyau est *suffisante*. Il n'est pas nécessaire de calculer le vecteur poids w pour spécifier la fonction de décision :

$$h(x) = \sum_{i=1}^m \alpha_i k(x_i, x'). \quad (4.4)$$

En un sens profond, une fonction noyau correspond à une mesure de similarité entre entrées x et x' , mesure qu'il est naturel d'utiliser en induction puisqu'un *a priori* évident est de supposer que deux entrées similaires doivent être associées à des sorties similaires. De fait, les

fonctions noyaux peuvent être considérées comme une *généralisation des fonctions de covariance*[106].

En effet, une fonction noyau peut s'exprimer comme une somme de produits sur un ensemble de fonctions :

$$k(x, x') = \sum_{i=1}^{\infty} \phi_i(x) \phi_i(x'). \quad (4.5)$$

4.2.3. Conditions pour avoir un noyau :

Il y a des conditions mathématiques, appelées théorème de Mercer [107], qui permettent de dire si une fonction est un noyau ou non, sans construire la projection dans l'espace des caractéristiques. En fait, il faut assurer que pour tout ensemble d'exemples de longueur l , la matrice $(k(\vec{x}_i), \vec{x}_j))_{1 \leq i, j \leq l}$, soit définie positive [108]. Nous ne nous attardons pas d'avantage sur ce point.

Exemples de noyaux :

- ✓ Noyau linéaire : $K(a, b) = \langle a, b \rangle$.
- ✓ Noyau polynomial : $K_k(a, b) = (1 + \langle a, b \rangle)^k$.
- ✓ Noyau RBF (gaussien): $K_a(a, b) = e^{-\frac{\|a-b\|^2}{2}}$.

4.2.4. Philosophie des méthodes à noyaux :

La section 2.1 a fourni une première illustration de l'utilisation de fonctions noyau comme moyen de transformer la recherche de régularités non-linéaires en recherche de régularités linéaires grâce au passage par un espace de redescription virtuel F (voir fonction 4.2). Quatre grande idées sous-tendent l'approche générale :

- Les données décrites dans l'espace d'entrée X sont projetées dans un espace vectoriel de redescription F
- Des régularités linéaires sont cherchées dans cet espace F .
- Les algorithmes de recherche n'ont pas besoin de connaître les coordonnées des projections des données dans F , mais seulement leurs produits scalaires.
- Ces produits scalaires peuvent être calculés efficacement grâce à l'utilisation de fonctions noyau.

Cette approche permet ainsi d'employer tout l'arsenal des méthodes linéaires développées jusque dans les années soixante pour découvrir des relations non-linéaires dans les données. Au fil des années récentes ont été ainsi revisités l'analyse en composantes principales (ACP)[112],

la méthode des filtres de Kalman[113], des méthodes linéaires de clustering[114], la discrimination linéaire de Fisher[115], etc.

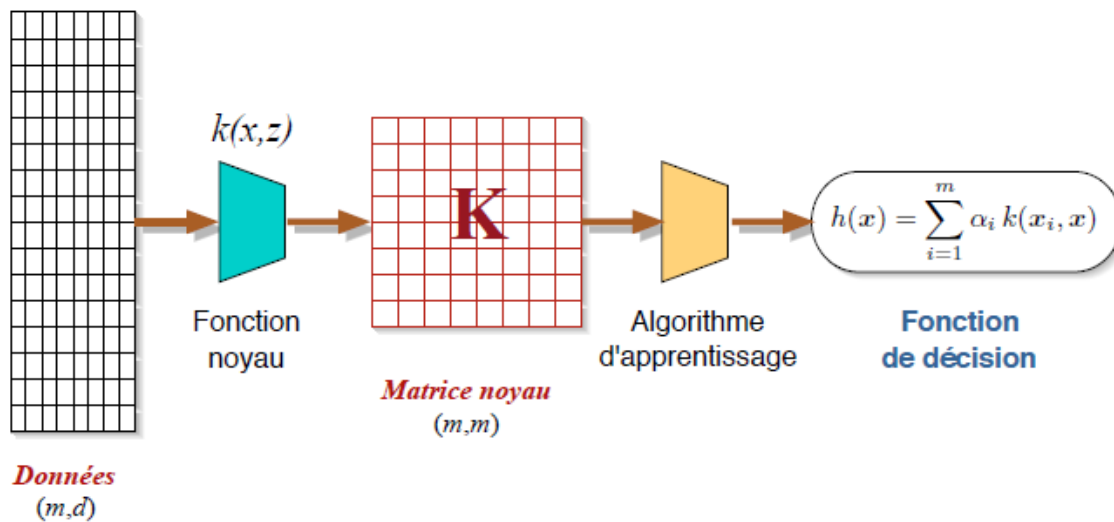


Fig.4.2. Chaîne de traitements générique des méthodes à noyaux.[116]

4.2.5. Coût et Régularisation :

Maintenant que nous avons présenté les noyaux, nous allons pouvoir préciser les deux critères minimisés lors de l'apprentissage : le risque empirique et la capacité. On parle alors de minimisation de coût régularisé.

La régularisation est une forme du contrôle de la capacité. L'idée est d'utiliser un terme de pénalisation.

Ce terme croît quand la fonction de décision augmente en complexité. Pour que la solution exacte du problème soit parcimonieuse il est nécessaire qu'au moins l'un de ces deux critères soit singulier (par opposition aux coûts réguliers) [117].

4.2.5.1. Définition - coût régulier :

Un coût est dit régulier s'il est infiniment dérivable.

C'est le cas par exemple du coût $C = (f(x_i) - y_i)^2$ utilisé pour la régression, du coût associé à la régression logistique $C = y_i \log f(x_i) + (1 - y_i) \log(1 - f(x_i))$ et des termes de pénalisation quadratiques.

Les coûts singuliers sont ceux qui conduisent à la parcimonie. Ils sont souvent construits à partir de valeurs absolues et sont singuliers à l'origine [117]. C'est le cas par exemple des coûts $C1 = |f(x_i) - y_i|$ et $C1(e) = \max(0, |f(x_i) - y_i| - e)$ utilisés pour la régression, du coût charnière $Ch = \max(0, y_i(f(x_i) + b - 1))$ utilisé en discrimination et des termes de pénalisation du type :

$$P = \sum_{i=1}^n |\alpha_i| \quad (4.6)$$

lorsque la solution f que l'on recherche vérifie :

$$f(x) = \sum_{i=1}^n \alpha_i k((,x)_b,x). \quad (4.7).$$

Le choix des fonctions de coût et de régularisation conduit à différents algorithmes (voir le tableau 4.1 pour quelques exemples). Les algorithmes minimisant un coût régulier sont de type Gaussien et sont très populaires pour leurs propriétés de dérivation et de facilité de calculs. Les algorithmes minimisant un terme singulier sont quant à eux très recherchés pour leur capacité à donner des solutions parcimonieuses. En revanche, l'utilisation de termes de type singulier impliquent des résolutions réputées lentes dans le pire des cas (on peut en particulier se référer au simplexe). Portnoy and Koenker [1997][118] montrent comment une méthode singulière peut être concurrentielle aux méthodes régulières d'un point de vue algorithmique. Nous allons montrer ici comment pratiquement faire mieux en tirant parti de la parcimonie.

Coût		Discrimination	Régression
sing.	sing.	LP SVM [Mangasarian,1998][137]	LP SVR : Ω_1
rég.	sing.	Regression logistique L_1	LARS : Ω_1 EFRON et al.2004.[139]
sing.	rég.	SVM : \mathcal{L}_h et Ω_2 :	SVR : Splines : Ω_2
rég.	rég.	k -régression logistique : \mathcal{L}_1 et Ω_2 . Lagrangian SVM : \mathcal{L}_r^2 [Mangasarian and musicant ,2001][138]	Splines : Ω_2

Tab 4.1. Récapitulatif de divers algorithmes d'apprentissage selon la nature de leurs fonctions objectifs.[119].

4.2.5.2. Signification et capacité d'approximation :

Dans ce qui a précédé on a montré que l'utilisation de fonctions noyau était liée à des raisons fondamentales diverses. En résumant, on peut dire que le choix d'une fonction noyau correspond implicitement au choix :

- d'une *mesure de similarité* entre éléments de l'espace d'entrée X ,
- d'une *projection* des éléments de X dans un espace vectoriel F doté d'un produit scalaire,
- d'un *espace fonctionnel* doté d'une base de fonctions analysantes (e.g. base de Mercer ou

base d'un espace de Hilbert à noyau reproduisant (RKHS)[120] dans laquelle est cherchée une fonction hypothèse, d'un *critère inductif régularisé* puisque le théorème de représentation montre que chaque noyau correspond à un terme de régularisation,

- d'une *fonction de covariance* définissant comment les éléments de X sont corrélés,
- d'une *mesure de probabilité sur un ensemble de fonctions* comme le montre l'interprétation des fonctions noyau comme fonction de covariance.

Le choix d'une fonction noyau est donc essentiel pour l'apprentissage et doit refléter au mieux toute connaissance *a priori* sur le domaine étudié.

4.3. Apprentissage à base de noyaux :

4.3.1. SVM :

4.3.1.1. L'idée fondatrice des SVM(séparateurs à vaste marge) [121] :

Ayant maintenant la ruse des noyaux à notre disposition, nous pouvons travailler, sans s'y projeter, dans des espaces de très grande dimension. Or une séparation linéaire, ainsi qu'une régression linéaire, est facilitée par la projection des données dans un espace de haute dimension... La contrepartie étant que la séparation qu'on y fait, facilement, ne signifie rien. Dit autrement, on a vite fait d'apprendre par cœur, c'est-à-dire d'apprendre une fonction qui ne se généralisera pas à de nouveaux exemples. Cette contrepartie est ce qu'on appelle la malédiction de la dimensionnalité. En maximisant la marge, les SVM s'en sortent malgré tout, et ne se font pas piéger par cette malédiction. En projetant dans l'espace des caractéristiques pour employer un algorithme de maximisation de la marge, on arrive à obtenir une séparabilité en conservant de bonnes capacités de généralisation, et c'est l'idée centrale des SVM.

4.3.1.2. L'hyperplan Séparateur Optimal [122] :

Le problème du classement binaire par un hyperplan consiste à trouver un hyperplan séparant l'espace des données en un sous-espace d'instances positives et un sous-espace d'instances négatives. Plus formellement, étant donné un ensemble de données

$$D = \{(x_1, y_1), \dots, (x_N, y_N)\}$$

avec $x \in \mathbb{R}^n$ et $y \in \{-1, +1\}$, le problème consiste à déterminer un hyperplan de vecteur normal

$$w \text{ tel que } : \forall (x_i, y_i) \in D : y_i(hw, x_i + b) \geq 0 \quad (4.7)$$

La solution y pour le vecteur x est alors donnée par : $y = \text{sign}(hw, x + b)$ [11] (4.8)

La minimisation du risque empirique, pour ce problème, conduit à un espace de solutions où toutes les hypothèses commettent le même nombre d'erreur sur D . La figure 4.3 illustre un cas avec plusieurs solutions.

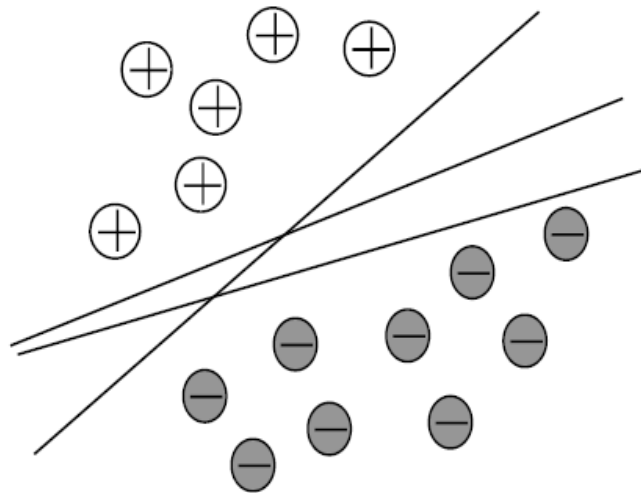


Fig.4.3.Exemples de séparateurs linéaires minimisant le risque empirique (cas séparable).[122]

4.3.1.3. Les SVM pour le classement :

Dans cette section, nous présentons les principaux algorithmes de classement basés sur les Séparateurs à Vaste Marge. Il est à noter que ces algorithmes peuvent aussi être utilisés pour d'autres objectifs que le classement. Par exemple, dans [123], une méthode itérative de sélection de variables a été proposée. Cette méthode repose sur des critères utilisant les paramètres des SVM.

4.3.1.4. Les SVM pour le cas séparables :

Dans cette section, nous décrivons la méthode pour déterminer l'hyperplan séparateur résultant de la minimisation du SRM dans le cas où les données sont linéairement séparables tel qu'illustré par la figure 4.4.

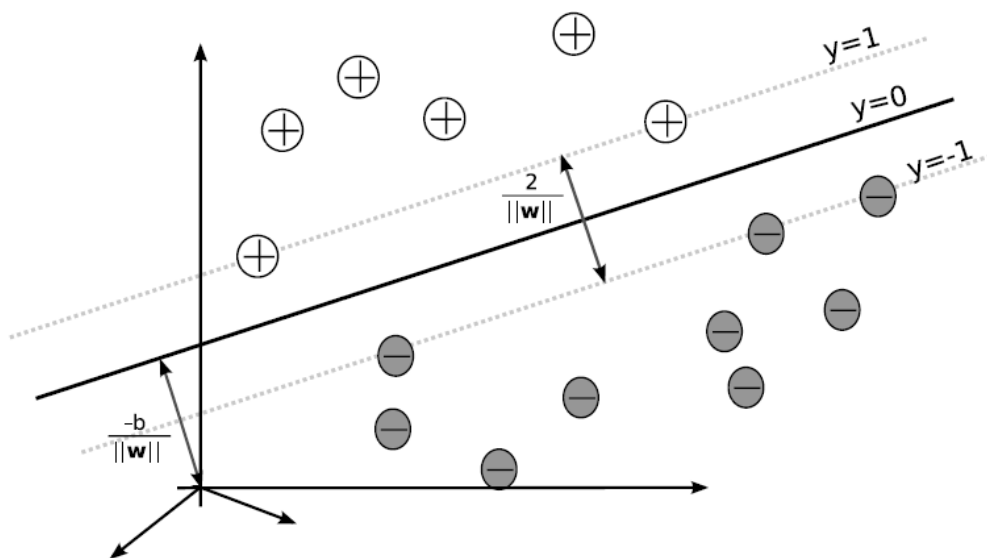


Fig.4.4.L'hyperplan séparateur dans le cas de données linéairement séparables.[124]

La maximisation de la marge, nous amène au problème quadratique suivant :

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 \\ \text{s. c.} \quad & y_i (\langle w, x_i \rangle + b) \geq 1 \\ & i \in \{1, \dots, N\} \end{aligned} \quad (4.9)$$

La formulation duale est en général plus utilisée pour résoudre le problème quadratique. En effet, la forme duale nous permet de passer d'un problème d'optimisation dans \mathbb{R}^n (dépendant de la dimension de l'espace) à un problème d'optimisation dans \mathbb{R}^N (dépendant du nombre d'exemples d'apprentissage). En utilisant, les multiplicateurs de Lagrange et les conditions de Karush-Kuhn-Tucker (KKT)[125].

4.3.2.5. Les C-SVM pour le cas non-séparables [126] :

Dans la majorité des cas, les données d'apprentissage, D , ne sont pas linéairement séparables. Pour répondre à ce problème, des variables d'écarts ξ_i sont introduites. Ces variables permettent d'autoriser, pour chaque point x_i , un certain degré d'erreur dans le classement :

$$y_i (\langle w, x_i \rangle) \geq 1 - \varepsilon_i \quad (4.10)$$

Le nombre d'erreur de classement permis est contrôlé par une valeur prédéfinie C . Cette valeur doit être fixée de manière empirique. L'introduction de C et des variables d'écarts nous amènent à un nouveau problème quadratique dans le primal :

$$\begin{aligned} \min_{w,b,\varepsilon} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \varepsilon_i \\ \text{s. c.} \quad & y_i (\langle w, x_i \rangle + b) \geq 1 - \varepsilon_i \\ & \varepsilon_i \geq 0 \\ & i \in \{1, \dots, N\} \end{aligned} \quad (4.11)$$

La formulation duale de ce problème est alors :

$$\begin{aligned} \min_{\alpha} \quad & -\sum_i \alpha_i + \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j \langle x_i, x_j \rangle \\ \text{s. c.} \quad & \sum_i y_i \alpha_i = 0 \\ & \forall i \in \{1, \dots, N\}, 0 \leq \alpha_i \leq C \end{aligned} \quad (4.12)$$

La solution est alors donnée par :

$$\begin{aligned} W &= \sum_{i=1}^N \alpha_i y_i x_i \\ B &= y_{usv} - \langle w, x_{usv} \rangle \end{aligned} \quad (4.13)$$

Il existe encore d'autres variantes des SVM, adaptées à des apprentissages bien particuliers, on cite :

- ✓ *v-SVM. [127]*
- ✓ *SVM Probabilistes. [128]*
- ✓ *Les SVM Semi-Supervisés. [129]*

4.3.2. Analyse en composante principale (PCA) [130] :

L'analyse en Composantes Principales (PCA) [] est un algorithme largement utilisé par de nombreuses applications telles que l'extraction de caractéristiques, la réduction de la dimensionnalité et la visualisation de données. Parmi ses avantages, sa facilité à calculer et à interpréter.

Etant donné x_1, \dots, x_n , le PCA est une projection orthogonale de ces points sur leurs axes principaux, qui sont ceux qui minimisent le coût de projection moyenne mesuré comme la distance au carré entre les points et leurs projections.

Le PCA résume l'algorithme jusqu'à une décomposition en valeurs propres de la matrice de covariance empirique des données $C_{emp} = E_{emp}[(x - E_{emp}(x))(x - E_{emp}(x))^T]$, i.e résoudre le système d'équations : $C_{emp}v_k = \lambda_k v_k$. Pour les données x_i de dimension d ce problème peut être résolu en un temps $O(d^3)$.

Dans [131] ce problème se pose dans l'espace de caractéristiques en remplaçant simplement x avec $\phi(x)$. Depuis la covariance empirique réside dans l'intervalle $\{\phi(x_1), \dots, \phi(x_n)\}$, pour simplifier la notation, nous supposons que nous avons déjà centrés les données dans l'espace des caractéristiques telles que : $\sum_{i=1}^N \varepsilon_i \phi(x) = 0$. Le problème de la valeur propre sera noté :

$$C_{emp}v_k = \frac{1}{n} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T v_k = \lambda_k v_k. \quad (4.14)$$

nous remarquons donc que les vecteurs propres sont de la forme : $v_k = \frac{1}{n} \sum_{i=1}^N \alpha_{ki} \phi(x_i)$.

En remplaçant ceci dans (2.58), nous trouvons que le problème des coefficients α sont faciles à calculer par le problème des valeurs propres : $K\alpha = \lambda\alpha$. (4.15)

K étant la matrice noyau de ces données. Après avoir résolu le problème des valeurs propres, nous pouvons calculer la projection d'un point x sur la $k^{\text{ième}}$ principale composante des données :

$$\langle v_k, \phi(x) \rangle = \sum_{i=1}^N \alpha_{ki} k(x, x_i).$$

Le noyau CPA peut être utilisé comme une étape de pré - traitement pour les algorithmes qui sont pas « kernelizable », ce sont les algorithmes qui ne reposent pas entièrement sur les produits scalaires .

4.3.3. Processus Gaussien [132] :

Il est possible de combiner l'approche de projection dans un espace de caractéristiques des méthodes à noyaux avec la modélisation et la règle de décision des méthodes génératives[133].

Etant donné les données d'apprentissage $\{(x_1; z_1), \dots, (x_n; z_n)\} \in E \times \{1, \dots, k\}$

sont des réalisations indépendantes d'un couple de vecteur aléatoire (X, Z) où z_j indique la classe de l'observation x_j . Ces données sont projetées dans un espace F à l'aide d'une fonction ϕ et le vecteur aléatoire $Y = \phi(X)$ est de plus supposé être, conditionnellement à $Z = i$, un processus gaussien de moyenne μ_i et d'opérateur de covariance Σ_i , $i = 1, \dots, k$.

La règle du maximum a posteriori (MAP) [134] peut être ensuite utilisée pour classer une nouvelle observation $\phi(x)$ dont la classe est inconnue. Malheureusement, si la fonction ϕ projette les données dans un espace de dimension infinie, il ne sera pas possible de construire avec une erreur raisonnable la règle de classification à partir d'un échantillon de taille finie.

Pour pallier ce problème, il a été proposé de contraindre la décomposition spectrale de l'opérateur de covariance de chaque classe C_i de sorte que celle-ci ait au plus d_i+1 valeurs propres différentes, $i = 1, \dots, k$ [135]. Les plus petites valeurs propres pour chaque classe sont donc supposées égales et de plus communes entre les classes. Cette modélisation peut également être contrainte pour donner naissance à 7 autre processus gaussiens parcimonieux. Si $d_i < n_i$ pour chaque classe où n_i est le nombre d'observations de la i ème classe, la règle du MAP peut être construite efficacement à partir d'un échantillon de taille finie. De plus, en exploitant le « kernel trick », on montre qu'il est possible de calculer la règle de classification du MAP au travers d'une fonction noyau sans avoir une connaissance explicite de la fonction ϕ .

Il est également possible d'étendre cette modélisation au contexte de la classification non supervisée (clustering) [136]. Dans ce cas, l'algorithme EM devra être utilisée car l'appartenance des données aux classes est inconnue (Z est donc considérée comme une variable manquante).

4.4. Conclusion :

Dans ce chapitre, nous avons introduit les concepts les plus fondamentaux des noyaux positifs et présenté des algorithmes qui s'appuient sur eux, qui sont capables de transformer tout algorithme linéaire, s'exprimant sous forme de produits scalaires, en un algorithme non linéaire.

Les méthodes à noyaux doivent leur nom à l'utilisation des fonctions du noyau, ce qui leur permet d'opérer dans l'espace des caractéristiques sans jamais calculer les coordonnées des données dans cet espace, mais plutôt par un simple calcul de produits scalaires entre les images de toutes les paires de données dans l'espace des fonctions. Cette opération est souvent moins coûteuse que le calcul explicite des coordonnées. Cette approche est appelée l'astuce du noyau (*kernel trick*). Les fonctions du noyau ont été introduites pour les données de séquence, des graphiques, du texte, des images, ainsi que des vecteurs.

Les noyaux positifs fournissent une mesure de similarité entre des objets complexes. Avec le cadre du risque régularisé, on peut mettre en œuvre la recherche sur les classes riches de fonctions et encore obtenir des fonctions qui peuvent être exprimées en une infinité de termes d'évaluations du noyau.

Beaucoup d'algorithmes fonctionnent sous ce principe mais, nous avons présenté uniquement trois en l'occurrence le svm, le PCA et processus gaussien.

CHAPITRE V

Classifieur à base de régression logistique à noyau.

5.1. Introduction :

Les récents développements technologiques et informatiques , parallèlement à la disponibilité de large quantité d'informations brutes ont contribué à l'émergence de nouvelles techniques et algorithmes dans le domaine de la reconnaissance et de la détection d'objet en général et des machines d'apprentissage, et la régression logistique à noyau en est un parfait exemple d'une large utilisation pour ces données [159].

Les principaux objectifs pour le développement de ces algorithmes comprennent l'identification des pattern à partir des données disponibles.

Comparée à d'autres méthodes à noyau , la régression logistique à noyau [140] est avantageuse en plusieurs points , à l'exemple elle génère des probabilités à posteriori en addition à la décision de classification et elle est capable de gérer un problème multi classe.

La KLR a déjà prouvé son efficacité dans plusieurs domaines, tels que tels que le diagnostic de cancer [141] et l'identification de speaker [142].

Dans ce chapitre nous allons présenter la régression logistique à noyaux et la mettre en œuvre pour la classification des images en pixels peau et non peau, sous Matlab.

5.2. Régression logistique :

5.2.1. Rappel régression logistique :

La régression logistique est un des modèles multi variables couramment utilisé en avec la régression linéaire multiple, la régression de Poisson et le modèle de Cox. Elle s'utilise lorsque la variable à expliquer (variable dépendante Y : survenue ou non d'un évènement) est qualitative, le plus souvent binaire. Les variables explicatives (variables indépendantes X_i) peuvent être par contre soit qualitatives, soit quantitatives. les variables indépendantes sont celles susceptibles d'influencer la survenue de cet événement c'est-à-dire les variables mesurant l'exposition à un facteur de risque ou à un facteur protecteur, ou variable représentant un facteur de confusion. Le développement de modèles pour données binaires a été favorisé par les besoins des biologistes [143]. En 1936, Fisher employait les termes d'analyse discriminante [144]. Le domaine médical s'est ensuite intéressé à ces techniques pour des études épidémiologiques, des pronostics de maladie, etc. [143]. Cox a joué un rôle fondamental dans la formulation du modèle de régression logistique, comme en témoigne, par exemple son ouvrage écrit en 1970 [145] ou celui écrit conjointement avec Snell [146]. Actuellement, c'est dans les domaines médical et

pharmaceutique que la régression logistique est la plus utilisée. L'emploi de la régression logistique est aussi présente dans d'autres domaines comme en économie, en agronomie et en sociologie [147].

5.2.2. Le modèle logit :

On supposera que la variable Y à laquelle on s'intéresse est la survenue ou non d'une maladie, dont les deux catégories seront notées M + et M-.

A une seule variable X explicative (équivalent d'une régression simple), le modèle s'écrit :

$$P(M+/X) = f(X) = \frac{e^{(\alpha + \beta x)}}{1 + e^{(\alpha + \beta x)}} \quad (5.1)$$

Il s'agit de la probabilité de maladie si la variable X est prise en compte et quand sa valeur est connue P (M + |X) se lit : probabilité de maladie si X. f (X) est la fonction logistique.

L'extension vers un modèle à plusieurs variables (régression multiple) se fait facilement :

$$P(M+/X_1 \dots X_n) = \frac{e^{(\alpha + \sum_{i=1}^n \beta_i x_i)}}{1 + e^{(\alpha + \sum_{i=1}^n \beta_i x_i)}} \quad (5.2)$$

5.2.3. Un cadre bayésien pour l'apprentissage supervisé :

Le classifieur bayésien est celui qui répond de manière optimale aux spécifications ci-dessus.

Pour un individu ω , il s'agit de calculer les probabilités conditionnelles (probabilité à posteriori) $P[Y(\omega)=y_k / X(\omega)]$ pour chaque modalité y.(se référer à la section 3.4)

5.2.4. Vraisemblance :

La vraisemblance (en Anglais *likelihood*) d'un échantillon Ω s'écrit:

$$L = \prod_{\omega} \pi(\omega)^{y(\omega)} \times (1 - \pi(\omega))^{(1-y(\omega))} \quad (5.3)$$

$$\text{Pour alléger l'écriture, on écrit : } L = \prod_{\omega} \pi^y \times (1 - \pi)^{(1-y)} \quad (5.4)$$

La vraisemblance correspond à la probabilité d'obtenir l'échantillon Ω à partir d'un tirage dans la population. Elle varie donc entre 0 et 1. La méthode du maximum de vraisemblance consiste à produire les paramètres $a = (a_0, a_1, \dots, a_J)$ de la régression logistique qui rendent maximum la probabilité d'observer cet échantillon [148] (page 81).

5.2.5. Log-vraisemblance :

Pour faciliter les manipulations, on préfère souvent travailler sur la log-vraisemblance notée LL (log-likelihood) qui est définie comme suit:

$$LL = \sum_{\omega} y \times \log \pi + (1 - y) \times \log (1 - \pi) \quad (5.5)$$

Le logarithme étant une fonction monotone, le vecteur a qui maximise la vraisemblance est le même que celui qui maximise la log-vraisemblance. Cette dernière en revanche varie entre $-\infty$ et 0 [149].

5.2.6. Déviance :

Comme alternative à la log-vraisemblance, bien souvent, on effectue la minimisation de la quantité:

$$D_M = -2 \cdot LL \quad (5.6)$$

appelée *déviance* [150] (page 13). Contrairement à la log-vraisemblance, elle est positive. L'objectif de l'algorithme d'optimisation est de minimiser cette déviance.

5.3. Régression logistique à noyaux

5.3.1 Astuce du noyau (Kernel Trick) :

Depuis de nombreuses années les méthodes à noyau ont démontré leur fort potentiel en discrimination ou régression, comme par exemple le classifieur *Support Vector Machine* [151], *kernel PCA* [152] ou bien encore *kernel Fisher discriminant* [153]. Pour plus de détails se référer à la section 4.2 du chapitre IV.

5.3.2. Régression logistique à noyau

La régression logistique à noyau est un modèle qui repose à la fois sur l'utilisation de la régression logistique et les méthodes à noyau. Sans passage par une sélection d'une partie des attributs des données (dans la plupart des techniques d'apprentissage) pour réduire la dimension de l'espace d'entrée, ce type de modèle impose que le nombre des paramètres soit linéairement lié au nombre des données d'apprentissage. Contrairement à la régression logistique classique, cette méthode est donc non paramétrique [154].

CH V. Classifieur à base de régression logistique à noyau.

La régression logistique à noyau est un modèle dont le log-ratio des probabilités conditionnelles est défini comme suit [154]:

$$\log \frac{P(Y = 1 / X)}{1 - P(Y = 1 / X)} = f(X) + b \quad (5.7)$$

Où f est une fonction appartenant à l'espace de projection à noyau (espace de Hilbert) H . Y est la variable binaire de sortie (variable endogène) et X correspond aux variables indépendantes de l'individu traité. La fonction f est définie comme suit [155]:

$$f(X) = \sum_{i=1}^n \alpha_i k(X, X_i) \quad (5.8)$$

Avec $\alpha_i = \frac{1}{\lambda} \left(\frac{y_i + 1}{2} - P(Y = 1 / x_i) \right)$ et k la fonction noyau.

Le modèle log-vraisemblance se présente sous la forme suivante [156]:

$$L = \sum_{i=1}^n \log \left(1 + e^{-y_i(f(x_i)+b)} \right) + \frac{\lambda}{2} \alpha^T K \alpha \quad (5.9)$$

où λ est un hyper paramètre, qui peut être ajusté par expérimentation, α le vecteur des paramètres α_i et K représente la matrice de Gram. L'ajustement optimal des paramètres α_i revient à minimiser l'équation 5.9 avec α_i différent de 0.

Une fois les paramètres α_i ajustés d'une manière optimale, la classification finale se fait selon l'équation suivante [154]:

$$P(y_i = 1 / X) = \frac{\exp \left(\sum_{i=1}^n \alpha_i k(X, X_i) + b \right)}{1 + \exp \left(\sum_{i=1}^n \alpha_i k(X, X_i) + b \right)} \quad (5.10)$$

Avec x l'individu à classifier et x_i le $i^{\text{ème}}$ individu de la base d'apprentissage.

5.4. La base d'image Compaq :

La construction de la base d'apprentissage est un élément important dans une démarche d'extraction de connaissances à partir des données. Pour un problème de classification de pixels en peau ou non, la qualité du corpus peut être jugée sur les facteurs suivants:

1. la taille de la base d'images et la variété dans les contenus des scènes, qui doit être représentative pour les différents sexes, races, et conditions d'éclairages.
2. la source de la base d'images, qui dépend particulièrement de l'application, tels que le filtrage de sites adultes sur le web et la détection de visage dans la vidéo.

Nous utilisons un corpus large d'images, contrairement à la majorité des travaux existants, qui pratiquent une phase d'apprentissage sur des classes prédéfinies d'images, sous des conditions d'éclairage connues à l'avance [61]. Si ces modèles conviennent généralement à des systèmes à base d'images spécifiques donc peu variées par exemple aux conditions d'éclairage, ils sont peu adaptés aux systèmes contenant une grande variété d'images.

Pour l'évaluation des performances des différents modèles de couleur de peau, des conditions de test doivent rester identiques. Malheureusement, il existe de nombreuses méthodes de détection de peau qui fournissent des résultats en se basant sur leurs propres bases d'images qui ne sont malheureusement pas accessibles. La base d'images la plus célèbre pour l'apprentissage et les tests est celle de Compaq [157].

Nous avons utilisé cette base d'image pour valider notre méthode de détection de la peau. Celle-ci inclut environ 4660 images contenant des pixels peau et 9000 ne contenant pas de pixels peau. Une vérité terrain de 4600 images binaires correspondant aux images contenant des pixels peau est disponible. Celui-ci est utilisé lors de la phase d'apprentissage pour identifier les parties peau de celles de non peau dans les images. Toutes les images de cette base sont obtenues à partir du Web. Différents types de peau appartenant à différentes races à savoir la race blanche, jaune et noire sont présentes sur ces images. La base d'images Compaq est caractérisée par l'hétérogénéité des conditions d'acquisition des images comme l'illumination variante, l'acquisition sous différents angles et positions, la diversité des appareils de capture des images avec différentes caractéristiques ainsi que la diversité des sources d'acquisition des images (capture à travers une vitre colorée ou non, un écran de télévision...). En outre, les personnes représentées sur ces images sont de différents âges (pour plus de détails voir la section 1.2 du premier chapitre). Cette base est donc sujette à toute sorte de distorsion, d'où toute la difficulté

CH V. Classifieur à base de régression logistique à noyau.

de classification des pixels issus des images lui appartenant en pixels peau et non peau. La figure 3.1 montre quelques exemples de ces images.



Fig. 5.1: Quelques images du corpus Compaq.

- Les images (1a) et (1b) sont brouillées.
- L'illumination est très variée dans les images (1c), (1d) et (2b).
- L'image (2d) est acquise à travers une fenêtre
- L'image (2a) est acquise à partir d'un écran de télévision.
- Le pilote sur l'image (2c) porte un casque coloré.

Toutes ces conditions créent des difficultés considérables pour la détection de la peau au sein de la base d'images Compaq.

5.5. Performances des techniques existantes :

5.5.1 Types de méthodes existantes :

CH V. Classifieur à base de régression logistique à noyau.

Tous les travaux rapportés sur la détection de la peau ont utilisé la classification comme stratégie de détection. Les techniques de détection basées sur la couleur peuvent être classifiées principalement en trois catégories, citées dans le chapitre III précédent, en l'occurrence :

Les méthodes explicites, les méthodes non paramétriques et les méthodes paramétriques.

5.5.2 Mesures d'évaluation :

Dans le cadre de l'évaluation de notre classifieur, nous considérons un problème simple de classification pour lequel nous nous intéressons à identifier l'appartenance de pixels issus de la base d'image (la classe peau ou la classe non peau). Nous indiquons par classe *positive* la classe peau et la classe *negative* la classe non peau. A cet effet, nous pouvons construire la *matrice de contingence* [**] (voir tableau 4.1) qui nous indique quatre types d'informations:

1)- le nombre de pixels peau classifiés correctement (nombre de *vrais pixels positifs NVP*), 2)-le nombre de pixels non peau classifiés correctement (nombre de *vrais pixels négatifs NVN*), 3)- le nombre de pixels peau mal classifiés (nombre de *faux pixels négatifs NFN*), 4)- le nombre de pixels non peau mal classifiés (nombre de *faux pixels positifs NFP*).

Cette matrice représente la répartition des pixels pour la tâche de classification binaire (peau et non peau). Les valeurs en gras correspondent aux documents correctement classés.

Classe trouvée \ Classe réelle	Classe peau	Classe non peau
	Classe peau	Classe non peau
Classe peau	NVP	NFP
Classe non peau	NFN	NVN

Tableau. 5.1: Matrice de contingence bi classe.

A partir de la matrice du tableau 5.1, nous définissons les mesures d'évaluation suivantes:

CH V. Classifieur à base de régression logistique à noyau.

Le *taux de vrais positifs (VP)* représente la probabilité qu'un pixel appartenant à la classe peau soit affecté à la classe peau. On emploie également souvent le terme de *rappel* pour désigner cette mesure.

Le *taux des faux positifs (FP)* correspond à la probabilité qu'un pixel appartenant à la classe non peau soit affecté à la classe peau.

Les équations dénotant ces mesures se présentent comme suit:

$$VP = \frac{NVP}{NVP + NFN} \quad (5.11)$$

$$FP = \frac{NFP}{NFP + NVN} \quad (5.12)$$

A partir de ces deux taux on peut déduire encore :

Le *taux des faux négatifs (FN)* qui correspond à la probabilité qu'un pixel appartenant à la classe peau soit affecté à la classe non peau.

Le *taux des vrais négatifs (VN)* qui correspond à la probabilité qu'un pixel appartenant à la classe non peau soit affecté à la classe non peau.

Les équations dénotant ces mesures se présentent comme suit:

$$FN = \frac{NFN}{NVP + NFN} \quad (5.13)$$

$$VN = \frac{NVN}{NVN + NFP} \quad (5.14)$$

Beaucoup de travaux effectués sur la détection de la peau en utilisant la base d'images Compaq pour la validation des méthodes qui leurs sont consacrées n'utilisent pas les mesures de FN et VN. Pour des fins de comparaison avec ces méthodes, nous utiliserons les mêmes critères d'évaluation et nous omettons donc nous aussi d'utiliser ces deux mesures. Durant

CH V. Classifieur à base de régression logistique à noyau.

l'expérimentation, plus le taux de VP est élevé et le taux de FP est bas, plus la méthode à évaluer est meilleure.

Classe réelle \ Classe trouvée	Classe réelle		Classe réelle \ Classe trouvée	Classe réelle	
	Classe peau	Classe non peau		Classe peau	Classe non peau
Classe peau	100	200	Classe peau	0	0
Classe non peau	0	0	Classe non peau	100	200
VP = 100%			VP = 0%		
FP = 100%			FP = 0%		

Tab. 5.2: Tous les pixels sont de la classe peau.

Tab. 5.3: Tous les pixels sont de la classe non peau.

Classe réelle \ Classe trouvée	Classe réelle		Classe réelle \ Classe trouvée	Classe réelle	
	Classe peau	Classe non peau		Classe peau	Classe non peau
Classe peau	100	0	Classe peau	0	200
Classe non peau	0	200	Classe non peau	100	0
VP = 100%			VP = 0%		
FP = 0%			FP = 100%		

Tab. 5.4: Classifieur parfait.

Tab. 5.5: Classifieur le pire.

Les deux premiers classifieurs représentent des classifieurs *radicaux* qui jugent soit tous les pixels comme étant peau, soit comme étant tous non peau. Les deux autres classifieurs sont des classifieurs *caricaturaux*, soit parfait (tous les pixels sont correctement classés), soit dramatique (les pixels sont toujours mal classés).

Dans [158], les auteurs ont utilisé la base d'image Compaq afin de valider le modèle qu'ils ont proposé. Ils ont utilisé d'autres critères d'évaluation que nous présentons comme suit:

- 1) *CDR* (*Correct Decision Rate*): pourcentage de pixels correctement classifiés parmi l'ensemble de tous les pixels;
- 2) *FAR* (*False Acceptance Rate*): pourcentage de pixels non peau mal classifiés parmi l'ensemble de tous les pixels;
- 3) *FRR* (*False Rejection Rate*): pourcentage de pixels peau mal classifiés parmi l'ensemble de tous les pixels.

A partir de la matrice du tableau 4.1, nous définissons les équations mathématiques dénotant ces mesures comme suit:

$$CDR = \frac{NVP + NVN}{NVP + NVN + NFP + NFN} \quad (5.15)$$

$$FAR = \frac{NFP}{NVP + NVN + NFP + NFN} \quad (5.16)$$

$$FRR = \frac{NFN}{NVP + NVN + NFP + NFN} \quad (5.17)$$

Plus *CDR* est élevé et *FAR* et *FRR* est petite, plus le système de classification de pixels est meilleur. Les meilleurs résultats obtenus dans [158] sont présentés dans le tableau 5.6.

5.5.3 Description du programme :

5.5.3.1. Etapes :

Le programme se déroule en deux phases : l'apprentissage et le test.

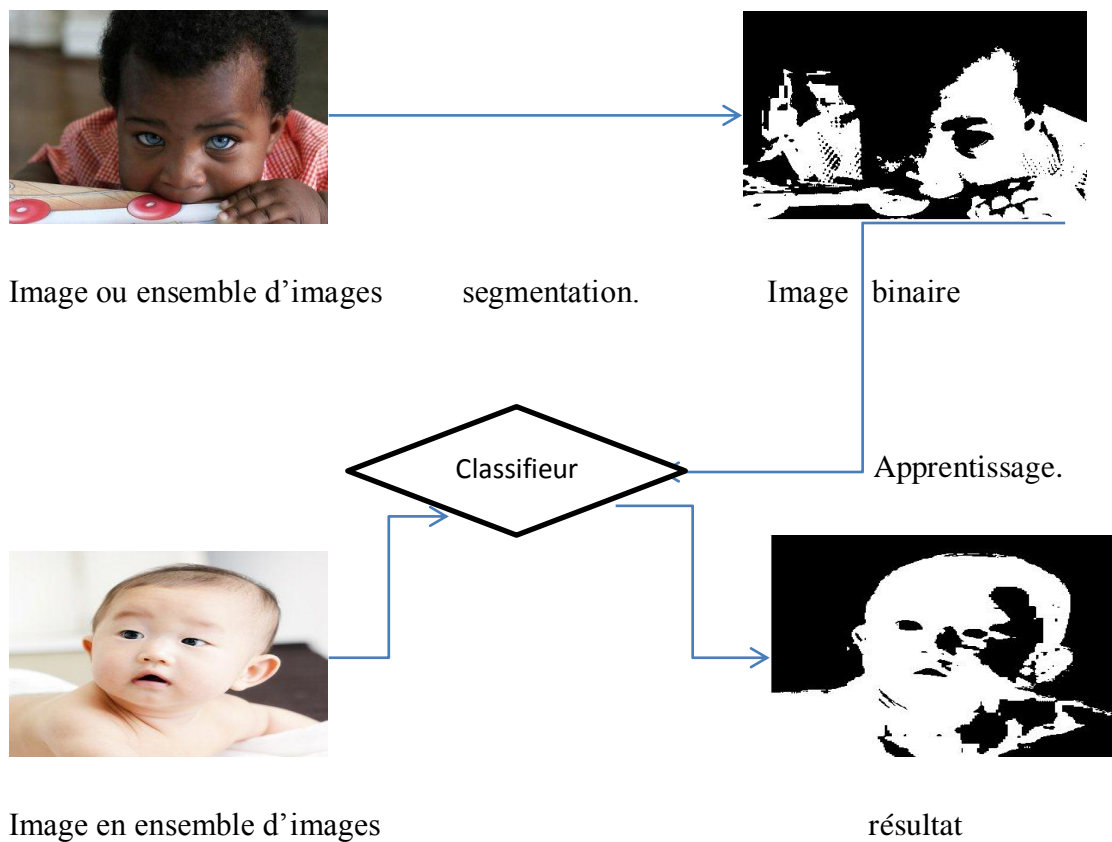


Fig 5.2. Déroulement du programme.

5.5.3.2. Test et résultats :

Nous allons présenter dans ce tableau quelques tests effectués dans différents espaces de couleur :

Les test sont déroulés de la manière qui suit :

Pour chaque groupe ethnique, on effectue un apprentissage dans chacun des trois espaces de couleurs sélectionnés pour le test et on effectue les test sur des tous les groupes ethniques.

Apprentissage Tests		Blanche			Noire			Jaune		
		Ycbr	HSV	NTSC	Ycbr	HSV	NTSC	Ycbr	HSV	NTSC
Blanche	VP	91.7836	84.7238	91.7486	98.1114	62.4363	98.1491	93.5943	63.401	93.5680
	FP	13.0254	20.7594	13.1453	25.7399	25.9552	26.5631	14.8738	22.577	16.2904
Noire	VP	53.4898	31.8494	53.4884	80.0757	50.5244	81.1885	58.1637	19.189	60.1354
	FP	19.3003	27.1118	19.3219	24.3253	26.5841	24.5143	19.9778	30.255	20.3218
Jaune	VP	77.9887	83.0516	77.7027	93.2402	49.6159	93.6083	80.4628	75.321	82.4065
	FP	17.5111	27.9903	17.5411	26.4302	30.1037	26.9778	18.5476	28.127	19.2303

Tab 5.6..Résultats sur différents espaces de couleurs.

Nous constatons que les meilleurs résultats sont enregistré dans les deux espaces ntsc et Y_cb.

5.5.4 Analyse comparative des techniques existantes :

Le tableau 5.6 illustre les méthodes ayant obtenues les meilleurs résultats sur la base d'images Compaq [157]. Toutes ces méthodes sont destinées à la détection de pixels peau appartenant à n'importe quelle race ethnique. Pour chaque méthode, on recense les références, l'espace de couleur utilisé pour la segmentation, l'intitulé de la méthode, le type de la méthode utilisée ainsi que l'ensemble d'apprentissage et de test utilisés ainsi que deux critères d'évaluation: le *taux de vrai positifs* et le *taux de faux positifs* (hormis la méthode proposée par les auteurs dans [158] qui

CH V. Classifieur à base de régression logistique à noyau.

utilisent les trois derniers critères d'évaluation que nous venons de présenter dans la section précédente). Pour plus de détails sur les types de ces méthodes, voir chapitre III.

A titre d'exemple, les auteurs Jones et Rehg [157] ont proposé deux méthodes basées sur l'apprentissage pour la détection de la peau; l'un est un modèle Bayésien non paramétrique basé sur les Histogrammes (Bayes SPM), l'autre est paramétrique et se base sur la mixture de gaussiennes. En utilisant la même stratégie d'apprentissage et de test, les auteurs opèrent deux évaluations pour chaque méthode. Les scores étant assez serrés, nous pouvons noter une légère supériorité de la méthode Bayes SPM.

<i>Ref.</i>	<i>Esp. coul</i>	<i>Comp .lumi n</i>	<i>Intitulé de la méthode</i>	<i>Type</i>	<i>Ens.Appr/ test</i>	<i>App- renti- ssage</i>	<i>VP</i>	<i>FP</i>
[157]	RGB	Oui	Bayes SPM	Non paramétrique	6822 /6818 images	Oui	80 90	8.5 14.2
	RGB	Oui	GMM	Paramétrique	6210/6818 images	Oui	90 80	15.5 9.5
[165]	TSL	Non	SOM.Map	Non paramétrique	n/a	Oui	78	32
[166]	RGB	Oui	Max.Ent. model	Paramétrique	2 000 000 000 / 2 000 000 000 pixels	Oui	82.9	10
[167]	Xyz	Non	Ellip.model	Paramétrique		Oui	90	20.9
	YCbCr	Non	SGM	Paramétrique	2000/6000	Oui	90	33.3
	YIQ	Non	GMM	Paramétrique	images	Oui	90	30
[168]	RGB	Oui	Bayes SPM	Non paramétrique	n/a	Oui	93.4	19.8
	YIQ	Oui	I-axis thresh	Explicite		Non	94.7	30.2
	RGB	Oui	Thresh.ratio	Explicite		Non	94.7	32.3
[169]	RGB	Oui	Beta mixt- ure models	Paramétrique	2 000 000/ 2 000 000 pixels	Oui	80 90 95	7.2 11.8 19.7
[170]	RGB	Oui	Impl.Math. model	Paramétrique	n/a	Oui	83.3 90.7	15.6 13.3
[163]	RGB	Oui	Color.Dist.	Explicite	4000/ 500	Non	89.97	9.26

CH V. Classifieur à base de régression logistique à noyau.

			Map		images				
[162]	RGB	Oui	Disc-DT based model	Non paramétrique	4167/9397 images	Oui	90	12.5	
[158]	Cb/Cr & Cr	Non	Comb.neu-ral networks	Non paramétrique	420 000/420 000 pixels	Oui	CDR	FAR	FRR
							82.61	14.66	2.73

Tab. 5.7: Comparaison des différentes méthodes de détection de la peau.

Les méthodes explicites proposées par Brand et al. [26] se justifient par la simplicité des règles de détection de la peau, la facilité de leur application et la faible complexité algorithmique. Un bon choix des intervalles de définition des pixels peau mènent à une performance de détection accrue. Cependant, nous remarquons que deux méthodes explicites sur trois possèdent un t taux de faux positifs important. La méthode *distance color map* de M. Abdullah-Al-Wadud et al. [163] apporte une nette amélioration sur le taux de faux positifs qui est beaucoup moins important que dans les modèles de Lee et al. [167]. Selon les scores obtenus, cette méthode surpasse toutes les autres en termes d'efficacité de détection. Le principal inconvénient de ce type de méthode est leur forte dépendance des conditions d'illumination. Les règles de seuillages ne sont efficacement valables que pour des conditions d'illumination précises. Un changement important sur ces conditions implique nécessairement de nouvelles règles à définir. En outre, la complexité à définir ces règles s'accroît si les pixels à traiter correspondent à différentes origines ethniques [171].

Les méthodes non paramétriques ont l'avantage d'être plus stables aux changements dans les conditions d'illumination variantes et la complexité de l'arrière plan et sont connus pour leur rapidité à la fois dans les phases d'apprentissage et de classification [160]. Le modèle Bayésien basé sur les Histogrammes (*Bayes SPM*) donne de meilleurs résultats que les modèles basés sur la distribution gaussiennes dans [157] ainsi que les modèles basés sur la définition de règles explicites dans [161] et démontre une bonne séparabilité entre les pixels peau et non peau quoique nous notons de faibles performances pour les modèles basés sur les réseaux de neurones de type *SOM (Self Organizing Map)*. L'inconvénient majeur de ces modèles est la nécessité d'avoir de larges collections de données pour l'apprentissage afin d'aboutir à de meilleures performances. En outre, ces méthodes sont coûteuses en termes d'espace mémoire. La meilleure méthode non paramétrique est celle de El Fkihi S. et al. [162]. Elle se classe parmi les meilleurs

CH V. Classifieur à base de régression logistique à noyau.

travaux avec des performance de détection assez proches que celles de M. Abdullah-Al-Wadud et al. [163] et de Zhanyu Ma et al. [164].

Les méthodes paramétriques permettent d'ajuster des distributions. La pertinence de ces méthodes et leur qualité de précision dépendent largement de la forme de distribution ainsi que l'espace de couleur choisi. Cependant, les phases d'apprentissage et de test pour ce type de méthodes sont lentes puisque celles-ci impliquent une procédure d'évaluation des paramètres utilisés. Parmi les avantages de ces méthodes, on peut citer le gain en espace mémoire ainsi qu'en possibilité de manipulation, plus d'intelligence dans la régularité des distributions et capacité d'interpoler les données d'apprentissage quand elles sont dispersées. La phase d'apprentissage peut être plus longue avec les approches paramétriques, mais celles-ci possèdent l'avantage de fournir une bonne estimation de la densité avec un ensemble d'apprentissage plus réduit [160]. Les meilleurs résultats offerts par ce type de méthodes est celui de Zhanyu Ma. [164]. Il représente un des meilleurs travaux en termes d'efficacité de détection avec celui de M. Abdullah-Al-Wadud et al. [163]. Cependant l'inconvénient que connaît ce type de méthodes (paramétrique) est la forte complexité de calcul algorithmique durant l'estimation des paramètres [164].

Bien que les méthodes présentées précédemment utilisent la base de données Compaq de manières différentes pour définir les images/ pixels d'apprentissage et les images/ pixels de test, et emploient différentes stratégies d'apprentissage, le tableau 4.6 donne une image assez fidèle des performances obtenues par ces méthodes. Une étude comparative est assez difficile à établir, mais néanmoins dans plusieurs travaux dont ceux cités sur le tableau précédent, des auteurs ont effectué une comparaison de leurs méthodes avec ceux de l'état de l'art malgré la diversité des stratégies d'apprentissage et de test utilisées [164] [162]. En analysant le tableau, on constate qu'il n'y a pas de type de méthode meilleur que les autres.

Notre travail est basé sur une méthode paramétrique pour la détection de tout type de peau appartenant à différents races ethniques. Il bénéficie donc de tous les avantages des méthodes paramétriques tout en assurant une complexité algorithmique raisonnable. La procédure d'évaluation des paramètres qui était excessivement lentes dans les autres méthodes paramétriques l'est moins dans notre cas. Comme toutes les autres méthodes citées dans le tableau 4.6, notre méthode est dédiée à la détection de tout type de peau appartenant à différentes races ethniques. Etant donné que l'efficacité d'une approche paramétrique est tributaire de l'espace de couleur choisi, nous avons choisi d'exploiter plusieurs espaces de couleur. Nous

présenterons par la suite les résultats de notre évaluation dans l'espace de couleur qui a donné les meilleurs résultats en comparaison avec ceux des méthodes présentées dans le tableau 5.6.

5.6. Conclusion :

Comme récapitulatif de ce chapitre :

nous avons présenté la méthode de régression logistique à noyau ,et une solution de classification des images en pixels peau et non peau ,qui est basée sur cette méthode.

Nous nous sommes servis de la base d'image Compaq pour la construction de notre corpus.

Nous avons effectué quelques test pour déterminer l'efficacité de notre approche, et ce après comparaison avec d'autres méthodes basées sur d'autres modèles de peau ainsi que sur des espaces de couleurs différents.

Les mesures d'évaluations principales sont les vrais positifs (faux positifs) qui correspondent au pixels peau (resp. non peau) classés comme pixels peau à la fin du test

Les meilleurs résultats ont été obtenus dans notre cas , sur l'espace Ycbcr, avec un taux de

93.5943 de VP et 14.8738% de faux positifs, les résultats aurait été meilleure sans la réduction du volume des groupe de test, et ce pour plus de rapidité de calcul.

CHAPITRE VI

Conclusions et perspectives.

A. Conclusion :

La détection de la peau humaine est un processus de recherche de pixels peau et non peau dans un image donnée ou vidéo.

Ce processus est souvent une étapes de prétraitement pour d'autre applications telles : la détection de visages, la vidéo surveillance ,la détection de nudité ... etc.

Un détecteur de peau de base transforme le pixel dans un espace de couleur de choix, et utilise un classifieur pour le labéliser , peu ou non peau. Le classifieur définit les règles de décisions de la classe de la couleur de la peau dans un espace de couleur et ce en se basant sur la base de données d'entraînement.

L'état de l'art montre qu'il existe dans la littérature de nombreuses méthodes pour la détection de régions de peau dans une image. Trois approches se détachent principalement : la détection basée sur l'extraction de traits caractéristiques du visage et /ou des mains, la détection basée sur une approche couleur, et la détection basée sur le mouvement. Cette dernière n'est jamais utilisée seule.

Les techniques basées sur l'extraction de traits caractéristiques nécessitent la définition de modèles pour ces traits (main, visage, corps humain...) qui sont associés à des templates. Le template matching consiste à comparer l'intensité de pixels entre un template prédéfini et plusieurs sous régions de l'image que l'on désire analyser. Cette méthode s'effectue par plusieurs balayages sur l'ensemble de l'image. Les endroits les plus propices à la présence de l'objet cherché sont identifiés par une différence minimale entre le template et l'image. Cette approche souffre de deux handicaps : la variation de formes d'un objet et une connaissance supposée de l'échelle de l'objet.

Les techniques basées sur le mouvement, telle que la méthode de différence d'image sont des techniques simples permettant de faire rapidement une estimation de la position d'un objet en mouvement. Cette estimation permet de réduire la zone de recherche d'un autre algorithme. On trouve cette approche souvent utilisée conjointement avec l'information de couleur de peau ou couplée avec un système basé sur la reconnaissance de traits caractéristiques.

Dans les méthodes basées sur la couleur, on distingue encore les méthodes paramétriques et les méthodes non paramétriques, ainsi que des méthodes basées sur une fixation empirique des règles et des seuils de décision. Pour les modèles paramétriques, leur pertinence et leur qualité de précision dépendent, dans une large mesure, de la forme de distribution ainsi que de l'espace de couleur choisi.

Quant aux méthodes basées sur la fixation empirique des règles et des seuils de décision, elles sont des méthodes simples et rapides ce qui explique leurs utilisations pour des applications qui

Conclusions & perspectives.

fonctionnent en temps réel. Néanmoins, on peut reprocher à ces méthodes la façon dont les règles de décision et les seuils sont choisis.

Les méthodes non paramétriques sont généralement rapides dans la phase d'apprentissage et de test. Elles ne font aucune hypothèse sur la répartition des données d'apprentissage et elles sont théoriquement indépendantes de la forme de distribution de la peau, contrairement aux méthodes paramétriques.

Contrairement aux phénomènes physiques issus de l'interaction entre la lumière et la matière, la mesure de la luminance et de la couleur fait intervenir la sensibilité du capteur. Tandis que la photométrie traduit d'un point de vue psychophysique la luminance reçue par un capteur, la colorimétrie s'intéresse à l'étude de la sensation colorée ainsi qu'à sa mesure objective, par l'utilisation des propriétés de trichromie. La quantité d'espaces couleur existants est si importante, du fait de la diversité des méthodes d'analyse couleur, qu'il paraît vain de vouloir en dresser une liste exhaustive. Néanmoins, nous en avons décrit quelques-uns, suffisamment représentatifs des différents types d'espaces rencontrés dans la littérature. On distingue notamment les espaces de primaires, les espaces uniformes, les espaces perceptuels, d'axes indépendants.

Pour la plupart de ces systèmes, leur création a été motivée par des applications.

Du fait de cette diversité, le choix de l'espace couleur le mieux approprié à une application n'est pas aisé et il n'existe malheureusement pas de loi universelle. Certains espaces, comme les systèmes hybrides, sont d'ailleurs dédiés à une seule et unique application.

Pour la plupart des études consacrées au choix du système colorimétrique dans différents cadres d'application (traitement d'images numériques [172], invariants couleur [173], quantification [174]) la conclusion générale montre que le choix d'un système couleur dans le cadre d'un traitement donné reste un problème largement ouvert. Même si quelques auteurs privilégient tel ou tel espace pour effectuer un type de traitement donné, il est aisé de trouver la contradiction dans un autre ouvrage.

Les espaces ($L^*a^*b^*$) et ($L^*u^*v^*$) sont à favoriser dès lors qu'il s'agit de mesurer des écarts de couleur, puisqu'ils permettent une meilleure relation entre distance couleur et écart colorimétrique. Néanmoins, pour être utilisés à bon escient, ils nécessitent la connaissance des conditions d'acquisition. Certains traitements nécessitent une décorrélation des composantes couleur. Dans ce cas, on privilégie généralement l'espace $L^*I^*I^*I^*$ de Ohta. Enfin, les systèmes de type Luminance Teinte Saturation sont intéressants pour définir des invariants couleur, étant données les propriétés d'invariance de la teinte vis-à-vis des ombres portées. Par contre, leur principal défaut est la présence de singularités pour des valeurs de saturations faibles.

Les algorithmes basés sur les noyau ont démontré à plusieurs reprises leur efficacité dans les tâches telles que les machines d'apprentissage, appliquées à la classification, la régression, ..etc.

Ils reflètent une bonne performance globale sur d'innombrables exemples de traitement de données, et tout cela en ajustant un nombre réduit de paramètres.

Ces méthodes se révèlent donc très adaptées aux problèmes d'apprentissage relatifs à la classification en reconnaissance d'images, avec de nombreuses qualités tels que la non-linéarité et la très grande dimension de l'espace. Mais elles ne sont pas sans inconvénients.

Problèmes des méthodes présentées : L'interprétation des résultats est souvent difficile à cause de la grande dimension. Il est difficile de tenir compte de la structure d'une image et de la position des objets analysés dans la construction d'un noyau pour images. Enfin, si la classification binaire se fait bien, la classification multi-classe reste un défi (on passe en général par plusieurs classificateurs binaires "un contre tous").

Nous avons adopté la régression logistique à noyau pour la construction de notre classifieur.

Les résultats obtenus sont bien meilleurs que beaucoup d'autres classifieurs, et ils se rapprochent de ceux des SVM, avec un avantage qu'elle renvoie des résultats probabilistes.

B. Perspectives :

Il n'est pas très probable qu'il ait une suite logique à ce travail, mais les différentes limites, personnelles ou portant sur certains travaux évoqués dans le manuscrit, risquent d'engendrer de nouveaux objectifs.

Il s'agit tout d'abord d'approfondir les connaissances liées aux systèmes de détection de la peau humaine sur les images et les vidéos pour recenser d'autres informations qui peuvent améliorer les performances de la technique de reconnaissance, ce qui nous vient à la tête est un détecteur de nudité sur des pages web, un travail qui a été réalisé par plusieurs personnes. Une astuce peut être, la fusion des résultats de classification. Cette dernière, peut assurer une reconnaissance efficace et robuste, et cela même dans des situations très complexes.

Il serait toutefois important d'intégrer l'information sonore et utiliser une nouvelle méthode de reconnaissance qui propose de détecter des séquences vidéo pour adultes.

Une autre information généralement associée aux documents vidéo est le texte. Ainsi, nous pouvons inclure cette information pour la reconnaissance des vidéos pour adulte.

Enfin, la méthodologie développée dans le cadre de ce travail peut être adaptée à d'autres problématiques.

Annexe A.

A / Modèle numérique anatomique de la peau .

La peau humaine est une enveloppe vivante de notre corps. L'anatomie de la peau humaine montre une structure multi-couches complexe qui est constituée de l'épiderme, du derme et de l'hypoderme. La complexité anatomique et physiologique de la peau humaine lui confère des propriétés complexes. Un essai sur l'avant bras gauche d'un homme, réalisé in vivo par la technique d'indentation couplée avec l'IRM, a donné des images de l'avant bras en deux états : non déformé et déformé sous l'indentation. Il en résulte, pour la première fois, une construction d'un modèle numérique anatomique d'une coupe de l'avant bras d'épaisseur de 16mm.

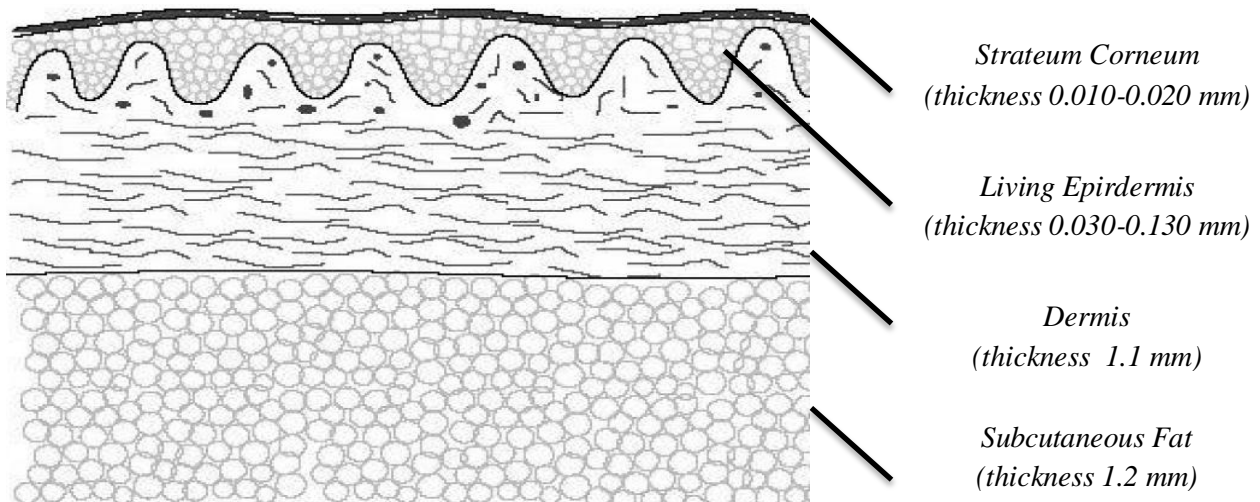


Fig A.1.L'Anatomie de la peau . HENDRIKS F. M., « Mechanical behaviour of human skin in Muscle /Bone are review, 2001.

B/ Physiologie de l'œil humain perception de la lumière :

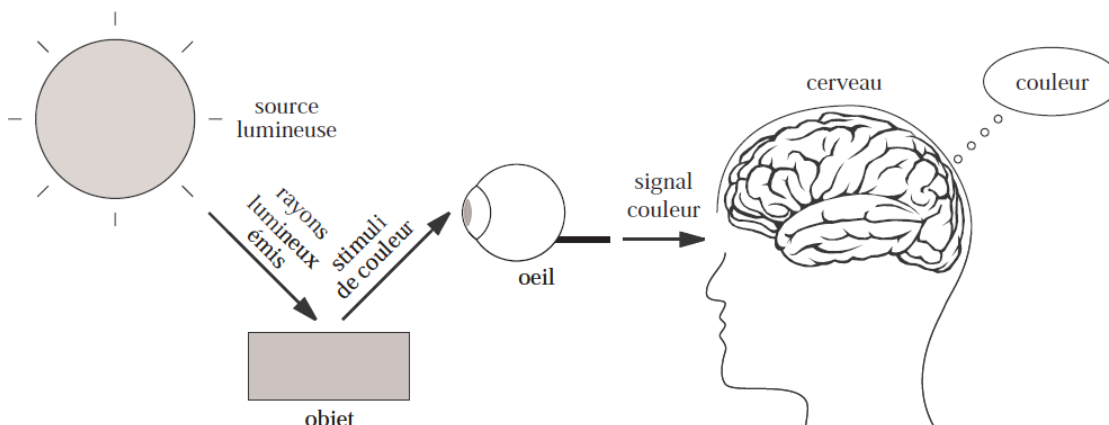


Fig.A.2.Perception humaine de la couleur.

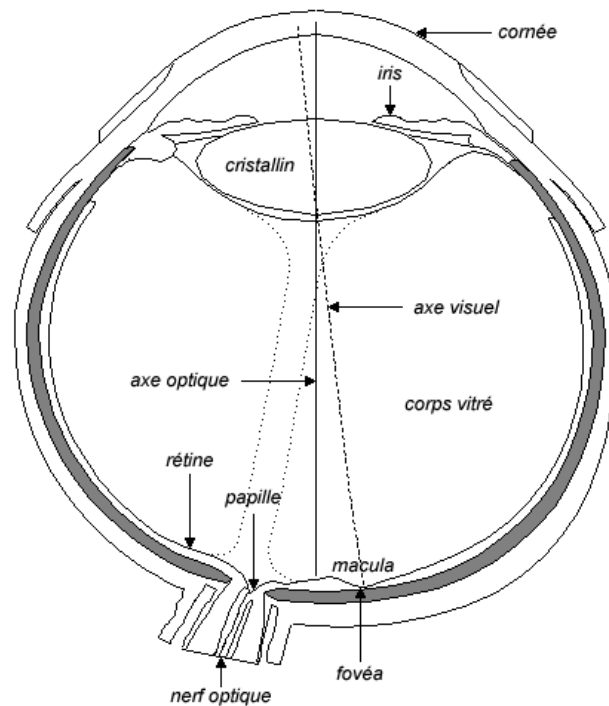


Fig.A.3. Coupe de l'œil humain. L'axe visuel passant par le centre du cristallin et celui de la fovéa est légèrement décalé vers le bas par rapport à l'axe optique qui passe par le centre du cristallin et celui de la pupille.

la rétine. Cette fonction est réalisée par les principaux éléments suivants :

– La **cornée** est une membrane transparente et résistante située sur la face avant de l'œil. Son rôle est de protéger l'avant du globe oculaire.

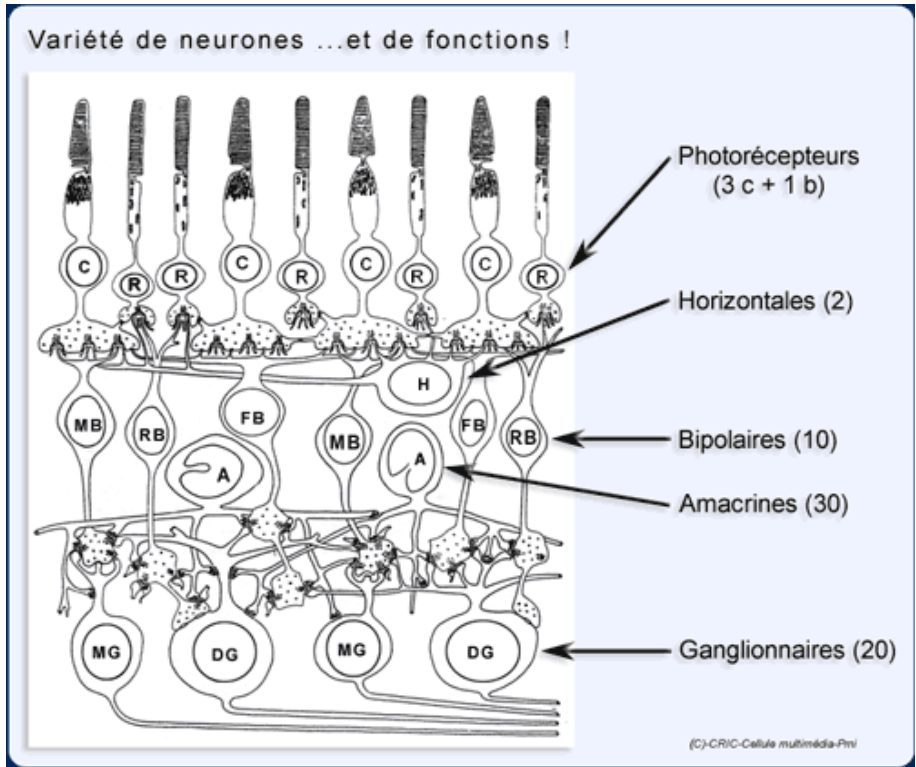
– **L'iris** est une membrane colorée qui fonctionne comme un diaphragme en contrôlant la quantité de lumière qui pénètre dans l'œil. Son ouverture centrale, la pupille est ainsi plus ou moins importante pour laisser passer plus ou moins de lumière.

– **Le cristallin** est une lentille biconvexe molle qui permet de focaliser le stimulus grâce à sa capacité à modifier sa courbure. Il nous permet d'avoir une vision nette de ce que nous observons.

– **Le corps vitré** est un liquide continuellement sécrété et absorbé, dont le rôle est d'assurer la structure autonome de l'œil.

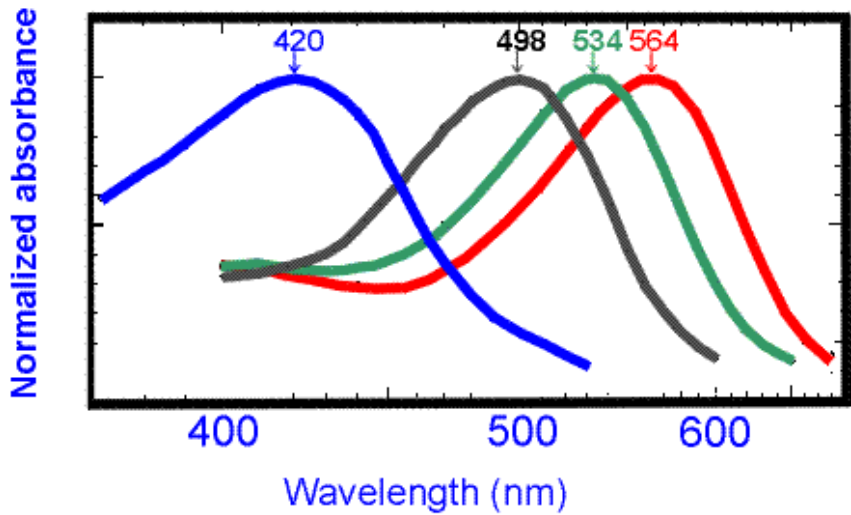
Après avoir traversé ces différents éléments, le stimulus couleur arrive sur une zone photosensible localisée au fond de l'œil, la rétine où viennent se projeter les images de la scène que nous observons. La rétine contient deux types de cellules photosensibles : les cônes et les bâtonnets.

Les bâtonnets permettent la vision nocturne (vision scotopique) tandis que les cônes permettent la vision diurne (vision photopique). Ces deux types de cellules n'ont pas la même sensibilité spectrale.



FigA.4 .Cônes et Bâtonnets.

- Cônes ($\sim 7 \cdot 10^6$) prépondérants dans la fovéa (2°) : **Vision diurne.**
- Bâtonnets ($\sim 120 \cdot 10^6$) prépondérants en périphérie : **Vision nocturne.**



Trois sortes de cônes :

- « R » sensibles aux grandes longueurs d'ondes, environ 30%
- « V » sensibles aux moyennes longueurs d'ondes, environ 60%
- « B » sensibles aux courtes longueurs d'ondes, environ 10%

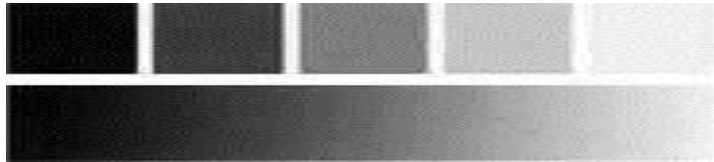
Annexe B.

A/ Lumière et couleur.

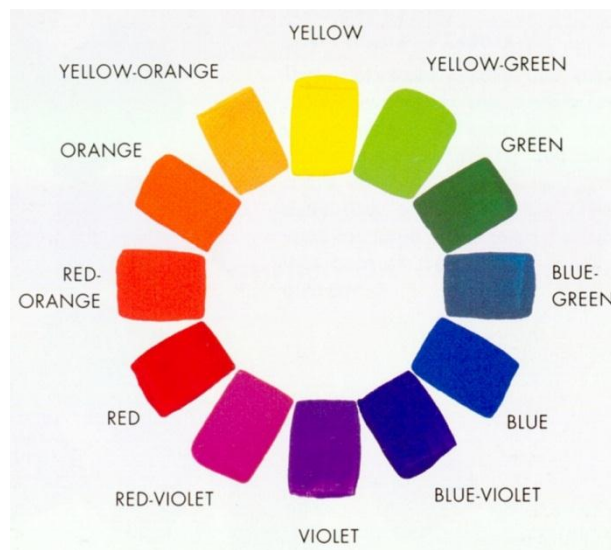
Quelques repères historiques

- Newton (1672) : spectre de la lumière blanche.
- Young (1801) , Helmholtz (1856) : trivariance de la couleur.
- Maxwell (1855) : fondateur de la colorimétrie.
- Munsell (1915) : Atlas avec équidistance perceptive.
- Guild et Wright : travaux de base de la CIE (1930).

Clarté



Teinte



Saturation



ANNEXE.

B /Atlas de Munsell :

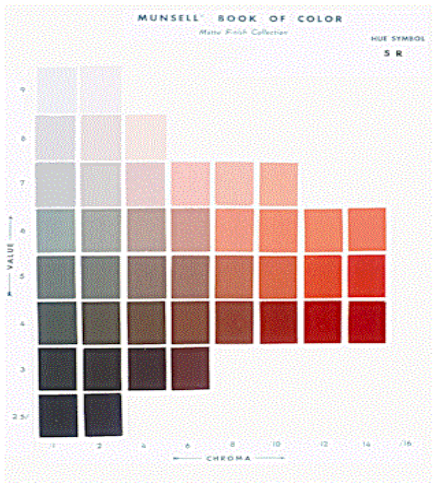
- A partir d'estimations « d'équidistance perceptive », subdivision fine de la roue des couleurs donnant les pages de l'atlas :

Une page = 1 teinte ou « Hue »

Actuellement : environ 1500 échantillons dans l'atlas

Un coloriste expérimenté distingue environ 10^6 couleurs !!

Clarté ou « value »

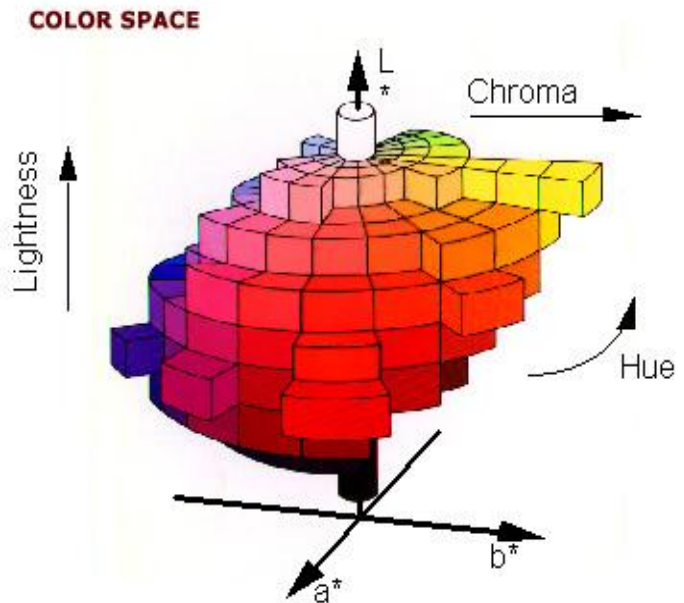


Saturation ou chroma

C /« Solide » des couleurs :

- Donne une représentation de l'espace des couleurs à laquelle se réfèrent peintres, coloristes...

- L Ordonnée verticale : clarté
- C distance à l'axe Oz : saturation
- h angle polaire : angle de teinte



C /Quelques Matrices de passage entre espaces de couleur..

Nous rappelons ici l'équation 1.20 permettant le changement de système de représentation de la couleur par une relation matricielle utilisant une matrice de passage P dans un sens et une matrice Q dans l'autre sens, avec $Q=P$.

$$\begin{bmatrix} R' * \\ G' * \\ B' * \end{bmatrix} = \begin{bmatrix} p11 & p12 & p13 \\ p21 & p22 & p23 \\ p31 & p32 & p33 \end{bmatrix} \times \begin{bmatrix} R * \\ G * \\ B * \end{bmatrix} = P \times \begin{bmatrix} R * \\ G * \\ B * \end{bmatrix}.$$

1-Transformation du système (R_c, G_c, B_c) avec illuminant E vers le système (X, Y, Z) :

$$P = \begin{bmatrix} 2.7690 & 1,7518 & 1,1300 \\ 1,000 & 4,5907 & 0,0601 \\ 0,0000 & 0,0565 & 5,5943 \end{bmatrix} = 5,6508 \times \begin{bmatrix} 0,490 & 0,310 & 0,200 \\ 0,177 & 0,812 & 0,011 \\ 0,000 & 0,010 & 0,990 \end{bmatrix} \quad [177],$$

2- Transformation du système (R_c, G_c, B_c) vers le système (Y', I', Q') :

$$P = \begin{bmatrix} 0,364 & 0,404 & 0,074 \\ 0,705 & -0,331 & -0,503 \\ 0,151 & -0,446 & 0,303 \end{bmatrix} \quad [176]$$

3- Transformation du système (R_E, G_E, B_E) avec illuminant E vers le système (Y', U', V') :

$$P = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,148 & -0,289 & 0,437 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} \quad [173].$$

4-3- Transformation du système (R, G, B) vers le système (I_1, I_2, I_3) ou système d'Ohta:

$$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix}$$

BIBLIOGRAPHIE.

- [1]. J. Yang, W. Lu, and A. Waibel, « Skin color modeling and adaptation », in proceeding of ACCV'98, Vol.2, pp. 687-694, Hong Kong, 1998.
- [2]. Q. Chen, H. Wu and M. Yachida, « face detection by fuzzy pattern matching », Proc. 5th Int. Conf. on Computer Vision, pp. 591-596, Cambridge, 1995.
- [3]. R-L. Hsu, M. Abdel-Mottaleb and A.K. Jain. Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), pp. 696-706, May 2002.
- [4]. V. Girondel, *Détection de peau, suivi de tête et de mains pour des applications multimédia*, DEA Signal Image Parole Télécom, Juillet 2002.
- [5]. H. Ouhaddi, P. Horain, K. Milkolajczyk. *Modélisation et suivi de la main*. Actes 4èmes Journées d'études et d'échanges Compression et REprésentation des Signaux Audiovisuels (CORESA'98), Lannion, France, 9-10 June 1998, pp. 109-114, 1998.
- [6]. M. M. Fleck, D. A. Forsyth and C. Bregler. Finding naked people. *European Conference on Computer Vision*, Springer-Verlag, Berlin, Germany, pp. 592-602, 1996.
- [7]. D. A. Forsyth and M. M. Fleck. Identifying nude pictures. *IEEE Workshop on the Applications of Computer Vision*, pp. 103-108, 1996.
- [8]. J. Z. Wang, J. Li, G. Wiederhold, and O. Firschein. Classifying objectionable websites based on image content. *Notes in Computer Science, Special issue on interactive distributed multimedia systems and telecommunication services*, pp.113-124, 1998.
- [9]. J. Z. Wang, J. Li, G. Wiederhold, and O. Firschein. System for screening objectionable images. *Images, Computer Communications Journal*, pp. 1355-1360, 1998.
- [10]. Navneet DALAL. 2006.Thesis Finding People in Images and Videos. 17 Jui 2006.
- [11]. L. Sigal, S. Sclaroff, and V. Athitsos. Skin color-based video segmentation under time-varying illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 862-877, July 2004.
- [12]. R. Brunelli and T. Poggio, *Face recognition: features versus templates*, *IEEE-T-PAMI*, Vol. 15, No. 10, pp.1042-1052, Octobre 1993.
- [13]. A. Pentland, B. Moghaddam and T. Starner, *View-based and modular eigenspaces for face recognition*, *IEEE-C-CVPR*, Seattle, WA, USA, pp. 84-91, 1994.
- [14]. J. C. Terrillon, M. Shirazi, H. Fumakachi and S. Akamatsu, *Comparative performance of different skin chrominance models and chrominances spaces for the automatic detection of human faces in color images*, in Proc. *IEEE International Conference on Face and Gesture Recognition*, Grenoble, pp. 54-61, March 2000.

BIBLIOGRAPHIE.

- [15]. C. C. Chiang, W. N. Tai, M. T. Yang, Y. T. Huang and C. J. Huang. *A novel method for detecting lips, eyes and faces in real time. Real-Time Imaging*, 9, 2003.
- [16]. V. Govindaraju, D. B. Sher, R.K. Srihari et S.N. Srihari, 'locating human faces in newspaper photographs', *Proc. Conf. On Comp. Vision and Pattgern Recognition*, pp. 549-554, 1989.
- [17]. A. Jacquin and A. Eleftheriadis, "Automatic location tracking of faces and facial features in video signal", *Int. Work. On Automatic Face and Gesture Recognition*, pp. 142-147, Zurich, 1995.
- [18]. S. Roux and E. Petit. 'Codeur H.263 amélioré par la visiophonie mobile, 7ème journée d'échange : Compression et représentation des signaux audiovisuels (CORESA), Dijon, Novembre 2001.
- [19]. J. B. Waite and W.J. Welsh, "an application of active contour models to head boundary location", *Proc. British Machine Vision Conf.*, pp. 407-412, Oxford, 1990.
- [20]. I. Craw, D. Tock et A. Bennett, " Finding face features", *Proc. 2nd European Conference on computer Vision*, pp. 92-96, 1992.
- [21]. T. F. Cootes, C. J. Taylor, D.H. Cooper and J. Graham, "Active shape models-their training and application", *comp. Vision and Image Understanding*, Vol. 61, No. 1, pp. 38-59, 1995
- [22]. T. Kanade, "Picture processing system by computer complex and recognition of humain faces", *Phd thesis, Departement of Information Science, Kyoto University*, Novembre 1973.
- [23]. G. J. Klinker, S.A. Shafer and T. Kanade, "A physical approach to color image understanding", *International Journal of Computer Vision*, Vol. 4, pp 7-38, 1990.
- [24]. Y. Sumi et Y. Ohta, "Detection of face orientation and facial components using distributed appearance modelling", *Proc. Int. Work. on Automatic Face and Gesture Recognition*, pp. 254-259, Zurich, 1995.
- [25]. A. Zelinsky and J. Heinzmann, "Real time visual recognition of facial gestures for human computer interaction", *2nd Int. Conf. On Automatic Face and Gesture Recognition*, pp. 351-356, Vermont, 1996.
- [26]. T. Leung, M. Burl et P. Perona, 'Finding faces in cluttered scenes using labelled random graph matching', *Proc. 5th Int. Conf. on Comp. Vision*, pp. 637-644, Cambridge, 1995.
- [27]. H. P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen et E. Patajan, "Multimodal system for locating heads and faces", *Proc.2nd Int. Conf. On Automatic Face and Gesture Recognition*, pp. 88-93, Vermont, 1996.

BIBLIOGRAPHIE.

- [28]. M. C. Burl and P. Perona, 'Recognition of planar object classes', *Comp. Vision and Pattern Recognition, San Fransisco, Juin 1996.*
- [29]. S. Roux and E. Petit. 'Codeur H.263 amélioré par la visiophonie mobile, 7ème journée d'échange : Compression et représentation des signaux audiovisuels (CORESA), Dijon, Novembre 2001.
- [30]. K. C Yow and R. Cipolla, "Finding initial estimates of humain face location", *Departement of Engineering, university of Cambridge, 1995.*
- [31]. A. Lemieux and M. Parizeau, *Flexible multi-classifier architecture for face recognition systems. Vision Interface, 2003.*
- [32]. C. R. Wren, A. Azarbayejani, T. Darrell and A. Pentland, *Pfinder : Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), pp. 780–785, 1997.*
- [33]. R. Cucchiara, C. Grana, M. Piccardi, A. Pratti and S. Sirotti: *Improving shadow suppression in moving object detection with hsv color information. Intelligent Transportation Systems, pp. 334– 339. IEEE, 2001.*
- [34]. P. L. Rosin and T. Ellis. *Image difference threshold strategies and shadow detection. In 6th British Machine Vision Conference, pp. 347–356, 1995.*
- [35]. A. Elgammal, D. Harwood and L. Davis, *Non-parametric model for background subtraction. European Conference on Computer Vision, pp. 751–767, 2000.*
- [36]. M. Alan, McIvor and *Background subtraction techniques. IVCNZ, 2000.*
- [37]. M. Talibi Alaoui, R. Touahni et A. Sbihi, *Classification des Images Couleurs par association des Transformations Morphologiques aux Cartes de Kohonen, CARI 2004, pp.83-90.*
- [38]. M. M. Fleck, D. A. Forsyth and C. Bregler. *Finding naked people. European Conference on Computer Vision, Springer-Verlag, Berlin, Germany, pp. 592-602, 1996.*
- [39]. D. A. Forsyth and M. M. Fleck. *Identifying nude pictures. IEEE Workshop on the Applications of Computer Vision, pp. 103-108, 1996.*
- [40]. R. Kjeldsen and J. kinder, "Finding skin in color images", *Proc. 2nd Int. Conf. on Automatic Face and Gesture Recognition, pp. 312-318, Vermont, 1996.*
- [41] R. Sève. *Physique de la couleur. De l'apparence colorée à la technique colorimétrique. Masson, 1996.*
- [42] A. Trouvé. *La mesure de la couleur : principes, technique et produits du marché. Afnor-Cetim, 1991.*
- [43] P. Kowaliski. *Vision et mesure de la couleur. Masson, seconde édition, 1990.*
- [44] Y. Dordet. *La colorimétrie : principes et applications. Eyrolles, 1990.*

BIBLIOGRAPHIE.

- [45] Association Française de NORmalisation. *Couleur, Colorimétrie*. AFNOR, 1989.
- [46] Y. Grall. *La fonction visuelle. Techniques de l'ingénieur, TE(E5175) :1–12, 11 1997.*
- [47] D. Dupont et D. Steen. *Colorimétrie : éléments théoriques. Techniques de l'ingénieur, R7(R6440) :1–, 2003.*
- [48] T. Carron. *Segmentations d'images couleur dans la base Teinte-Luminance-Saturatuion : approche numérique et symbolique. Thèse de doctorat, Université de Savoie, décembre 1995.*
- [49] N. Vandenbroucke. *Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football. Thèse de doctorat, Université des Sciences et Technologies de Lille, décembre 2000.*
- [50] D. L. MacAdam. *Color measurement, theme and variation. Optical Sciences. Springer-Verlag, second revised édition, 1985.*
- [51] D. B. Judd. *A Maxwell triangle yielding uniform chromaticity scales. Journal of the Optical Society of America, 25 :24–35, 1935.*
- [52] L. Bedat. *Aspects psychovisuels de la perception des couleurs. Application au codage d'images couleur fixes avec compression de l'information. Thèse de doctorat, Université de Nantes, octobre 1998.*
- [53] O. D. Faugeras. *Digital color image processing and psychophysics within the framework of a human visual model. Ph. D. dissertation, University of Utah, juin 1976.*
- [54] C. Garbay. *Modélisation de la couleur dans le cadre de l'analyse d'images et de son application à la cytologie automatique. Thèse de doctorat, Institut National Polytechnique de Grenoble, décembre 1979.*
- [55] D. H. Ballard et C. M. Brown. *Computer vision. Prentice-Hall, New Jersey, 1982.*
- [56] V. Rodrigues. *L'analyse du mouvement dans les images couleurs. Thèse de doctorat d'état, Université Paul Sabatier de Toulouse, novembre 1985.*
- [57] L. Brun. *Segmentation d'images couleur à base topologique. Thèse de doctorat, Université de Bordeaux I, décembre 1996.*
- [58]. L. Jordao, M. Perrone, J. P. Costeira, and J. Santos-Victor. *Active face and feature tracking. In Proceedings of International Conference on Image Analysis and Processing, Venice, Italy, pp. 572-576, September 1999.*
- [59]. D. Saxe and R. Foulds. *Toward robust skin identification in video images. 2nd International Face and Gesture Recognition Conference, Septembre 1996.*
- [60]. R. Kjeldsen and J. kinder, "Finding skin in color images", *Proc. 2nd Int. Conf. on Automatic Face and Gesture Recognition, pp. 312-318, Vermont, 1996.*

BIBLIOGRAPHIE.

- [61]. M. Hammami, "Modèle de peau et application à la classification d'images et au filtrage des sites Web". Thèse de doctorat en informatique à l'école nationale de Lyon, 2005.
- [62]. R. O. Duda, P. E. Hart. « *Pattern classification and scene analysis* » John Wiley, 1973.
- [63] G. Saporta. *Probabilités, Analyse de données et Statistique*. Technip, 1990.
- [64]. E. Parzen, "On estimation of a probability density function and mode", *Ann. Math. Stat.*, Vol.33, pp. 1065-1076, 1962.
- [65]. R.O. Duda and P.E. Hart, "Pattern classification and scene analysis", John Wiley, 1973.
- [66]. R-L. Hsu, M. Abdel-Mottaleb and A.K. Jain. *Face detection in color images*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), pp. 696-706, May 2002.
- [67]. E. Saber and A. Tekalp. *Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions*. *Pattern Recognition Letters*, 19(8), pp. 669-680, 1998.
- [68]. M. Yang and N. Ahuja. *Detecting human faces in color images*. *International Conference on Image Processing (ICIP)*, Chicago, pp.127– 130, 1998.
- [69]. M. J. Jones and J. M. Rehg. *Statistical color models with application to skin detection*. *Technical Report Cambridge Research Laboratory, CRL 98/11, Compaq*, 1998.
- [70]. J. C. Terrillon, M. N. Shirazi, H. Fukamachi, and S. Akamatsu. *Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images*. In *Fourth International Conference on Automatic Face and gesture Recognition*, pp. 54-61, 2000.
- [71]. M. R. Teague, *Image analysis via the general theory of moments*. *Journal of the Optical Society of America*, 70(8), pp. 920–930, 1980.
- [72]. S. Marcel, O. Bernier et D. Collobert, *Approche EM pour la construction de régions de teinte homogènes : application au suivi du visage et des mains d'une personne*, Université de Poitiers, CORESA 2000, Octobre 2000.
- [73]. M. Hunke and A. Waibel, "Face locating and tracking for humain computer interaction", *IEEE Computer*, pp. 1277-1281, November 1994.
- [74]. J. Yang, W. Lu, and A. Waibel, « *Skin color modeling and adaptation* », in *proceeding of ACCV'98, Vol.2*, pp. 687-694, Hong Kong, 1998.

BIBLIOGRAPHIE.

- [75]. J. Y. Lee and S. I. Yoo. *An elliptical boundary model for skin color detection*. In *Proc. International Conference on Imaging Science, Systems and Technology, Las Vegas, USA, June 2002*.
- [76]. B. D. Zarit, B.J. Super, and F. K. H. Quek. *Comparison of five color models in skin pixel classification*. In *Proceedings. International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Corfu, Greece, pp.58-63, September 1999*.
- [77]. Q. Chen, H. Wu and M. Yachida, « *face detection by fuzzy pattern matching* », *Proc. 5th Int. Conf. on Computer Vision*, pp. 591-596, Cambridge, 1995.
- [78]. G. Gomez. *On selecting colour components for skin detection*. In *Proceedings. 16th International Conference on Pattern Recognition*, vol. 2, pp. 961-964, 2002.
- [79]. R. Schumeyer, and K. Barner. *A color-based classifier for region identification in video*. In *Visual Communications and Image Processing 1998, SPIE*, vol. 3309, pp. 189–200, 1998.
- [80]. L. Sigal, S. Sclaroff, and V. Athitsos. *Estimation and prediction of evolving color distributions for skin segmentation under varying illumination*. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 152–159, 2000.
- [81]. M. Soriano, B. Martinkauppi, S. Huovinen, and M. Laaksonen. *Skin detection in video under changing illumination conditions*. In *Proc. Computer Vision and Pattern Recognition*, vol. 1, pp. 839-842, 2000.
- [82]. S. Birchfield. *Elliptical head tracking using intensity gradients and color histograms*. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Santa Barbara, CA*, pp. 232-237, 1998.
- [82]. A. Wu, M. Shah, and N. da Vitoria Lobo. *A virtual 3d blackboard: 3d finger tracking using a single camera*. In *Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)*, March 2000.
- [83]. F. K. H. Quek, T. Mysliwiec, and M. Zhao. *Fingermouse: A freehand pointing interface*. *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition, Zurich, Switzerland*, pp. 372-377, June 1995.
- [83]. M. J. Jones, and J. M. Rehg. *Statistical color models with application to skin detection*. In *Proc. of the CVPR '99*, vol. 1, pp. 274–280, 1999.
- [84]. D. Saxe and R. Foulds. *Toward robust skin identification in video images*. *2nd International Face and Gesture Recognition Conference, Septembre 1996*.
- [85]. M. J. Swain and D.H. Ballard, « *Color Indexing* ». *International Journal of Computer Vision*, 1998.

BIBLIOGRAPHIE.

- [86]. M. J. Jones and J. M. Rehg. *Statistical color models with application to skin detection. International Journal of Computer Vision*, 46(1), pp. 81-96, January 2002.
- [87]. D. Chai and A. Bouzerdoun. *A bayesian approach to skin color classification in YCbCr color space. In Proc. IEEE Region Ten Conference (TENCON'2000), Vol. 2, pp. 421-424, 2000.*
- [88]T. Kohonen, "*Self-Organizing Maps*". Springer Verlag, 1995.
- [89]D. Brown, I. Craw, J. Lewthwaite. *A SOM based approach to skin detection with application in real time systems, BMVC01, 2001.*
- [90]M. J. Jones, and J. M. Rehg. *Statistical color models with application to skin detection. In Proc. of the CVPR '99, vol. 1, 274-280, 1999.*
- [91]P. Kakumanu, S. Makrogiannis, R. Bryll, S. Panchanathan, and N.
- [92]Bourbakis. *Image chromatic adaptation using ANNs for skin color adaptation, Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence, ICTAI04.*
- [93]P. Kakumanu. *A face detection and facial expression recognition method applicable to assistive technologies and biometrics, PhD Dissertaion, CSE Department, Wright State University, 2006.*
- [94]. P. Peer, J. Kovac, and F. Solina. *Human skin colour clustering for face detection. International Conference on Computer as a Tool, EUROCON 2003, Ljubljana, Slovenia, September 2003.*
- [95]. A. Trémeau et al. *Image couleur : de l'acquisition au traitement. Collection Sciences Sup, 480 pages, dunod éditions, ISBN 2 10 006843 1, 2004.*
- [96]. H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang. *Color image segmentation: advances and prospects. Pattern Recognition*, 34(12), pp. 2259-2281, December 2001.
- [97]. N. Mottin, *Localisation de visages dans des images acquises avec différents cadrages de caméras. Application à l'indexation, DEA ARAVIS (image et vision) Université Sophia Antipolis, Juin 2000.*
- [98]. A. Trémeau et al. *Image couleur : de l'acquisition au traitement. Collection Sciences Sup, 480 pages, dunod éditions, ISBN 2 10 006843 1, 2004.*
- [99]. N. Mottin, *Localisation de visages dans des images acquises avec différents cadrages de caméras. Application à l'indexation, DEA ARAVIS (image et vision) Université Sophia Antipolis, Juin 2000.*
- [100] Marc Atteia et Jean Gaches , *Approximation Hilbertienne - Splines, Ondelettes, Fractales, PUG, 1999*
- [101] V. N. Vapnik. *Statistical learning theory. Wiley, New York, (1998).*

BIBLIOGRAPHIE.

- [102] *Machines à Vecteurs Supports .Didacticiel-Hervé Frezza-Buet, Supélec.2012*
- [103] S. Mika, G. Ratsch, J. Weston, B. SchPolkopf, K.-R. Muller, *Fisher discriminant analysis with kernels, Proceedings of IEEE Neural Networks for Signal Processing Workshop, 1999,*
- [104] J. Weston, A. Gammerman, M. Stitson, V. Vapnik, V. Vovk, C. Watkins, *Support vector density estimation, in: B. SchPolkopf, C. Burges, A. Smola (Eds.), Advances in Kernel Methods: Support Vector Machines, MIT Press, Cambridge, MA,*
- [105] N. Cristianini, C. Campbell, C. Burges (Eds.), *Support vector machines and kernel methods, Machine Learning, 2001, to appear.*
- [106] A. Jain and D. Zongker. *Feature selection : Evaluation, application and small sample performance. IEEE Transaction on Pattern Analysis and Machine Intelligence, 19(2) :153-158, (1997).*
- [107] J. Mercer, *Functions of positive and negative type and their connection with the theory of integral equations, Philos. Trans. Roy. Soc. London A 209 (1909) 415–446.*
- [108] J. Mercer, *Functions of positive and negative type and their connection with the theory of integral equations, Philos. Trans. Roy. Soc. London A 209 (1909) 415–446.*
- [109] A. Luntz and V. Brailovsky. *On estimation of characters obtained in statistical procedure of recognition. Technicheskaya Kibernetica, 3, (1969).*
- [110] Trevor Hastie, Robert Tibshirani and Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2001*
- [111] Olivier Chapelle, Bernhard Scholkopf and Alexander Zien, *Semi-supervised Learning, MIT press 2006*
- [112] *Kernel Principal Components Analysis. MaxWelling Department of Computer Science University of Toronto 10 King's College Road Toronto, M5S 3G5 Canada welling@cs.toronto.edu*
- [113] Alain Berline and Christine Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics, Kluwer Academic Publishers, 2003*
- [114] M. Soriano, B. Martinkauppi, S. Huovinen, and M. Laaksonen, "Adaptive skin color modeling using the skin locus for selecting training pixels," *Pattern Recognition, vol. 36, no. 3,*
- [115] John Shawe-Taylor and Nello Cristianini *Kernel Methods for Pattern Analysis, Cambridge University Press, 2004*
- [116] Léon Bottou, Olivier Chapelle, Dennis DeCoste and Jason Weston *Large-Scale Kernel Machines (Neural Information Processing, MIT press 2007*
- [117] Mila Nikolova. *Local strong homogeneity of a regularized estimator. SIAM Journal on Applied Mathematics, 61*

BIBLIOGRAPHIE.

(2) :633–658, 2000.

[118] Stephen Portnoy and Roger Koenker. *The gaussian hare and the laplacian tortoise : computability of squared-error*

versus absolute-error estimators. Statistical Sciences, 1997.

[119] M. Y. Park and T. Hastie. *L1 Regularization Path Algorithm for Generalized Linear Models. Technical report, Stanford University, February (2006).*

[120] Halmos, P.R. *A Hilbert Space Problem Book. D. Van Nostrand Company, Inc., 1967.*

[121] Cortes, C. & Vapnik, V. (1995), 'Support-vector networks', *Machine Learning* 20(3), 273–297.

[122] V. N. Vapnik. *The Nature of Statistical Learning Theory. Springer Verlag, New York, (1995).*

[123] David, A. & Lerner, B. (2005), 'Support vector machine-based image classification for genetic syndrome diagnosis', *Pattern Recognition Letters* 26, 1029–1038

[124] J. Neumann, C. Schnörr, and G. Steidl. *Combined SVM-Based Feature Selection and Classification. Machine Learning, 61(1-3) : 129-150, (2005).*

[125] J. C. Platt, N. Cristianini, and J. Shawe-Taylor. *Large margin DAGs for multiclass classification. In S. A. Solla, T. K. Leen, and K. R. Müller, editors, Advances in Neural Information Processing Systems 12, pages 547-553, The MIT Press, (2000).*

[126] Aizerman, M.A., Braverman, E.M. and Rozoner, L.I. *Theoretical foundations of the potential function method in pattern recognition learning. Automation and Remote Control, 25:821–837, 1964.*

[127] Shevade, S. K., Keerthi, S. S., Bhattacharyya, C., and Murthy, K. R. K. (1999). *Improvements to smo algorithm for svm regression. Technical Report CD-99-16, National University of Singapore.*

[128] Platt, J. (1998). *Fast training of support vector machines using sequential minimal optimization. In Schölkopf, B., Burges, C., and Smola, A., editors, Advances in Kernel Methods : Support Vector Machines. MIT Press, Cambridge, MA.*

[129] J. H. Friedman. *Another approach to Polychotomous classification. Technical report, Department of Statistics, Stanford University, (1996).*

[130] [Aro50] Aronszajn, N. *Theory of reproducing kernels. Trans. Amer. Math. Soc., vol. 68:337{404, 1950.*

[131] Schölkopf, B., Smola, A. J., and Müller, K.-R. *Nonlinear component analysis as a kernel Eigenvalue problem. Neural Comput., vol. 10:1299{1319, 1998.*

[132] B. Schölkopf and A. J. Smola. *Learning with Kernels: Support*

BIBLIOGRAPHIE.

- Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.
- [133] PG Equipe MISTIS, INRIA Rhône-Alpes & LJK stephane.girard@inria.fr.
- [134] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [135] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge Univ. Press, 2003.
- [136] Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press.
- [137] O. Mangasarian. *Generalized support vector machines*. NIPS Workshop on Large Margin Classifiers, 1998. URL citeseer.ist.psu.edu/mangasarian98generalized.html.
- [138] O. L. Mangasarian and David R. Musicant. *Lagrangian support vector machines*. *Journal of Machine Learning Research*, 1 :161–177, 2001.
- [139] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. *Least angle regression*. *Annals of Statistics*, 32(2) :407–499, 2004. URL citeseer.ist.psu.edu/efron02least.html.
- [140] Jaakkola, T. S., and Haussler, D. 1999. *Probabilistic kernel regression models*. In *Proceedings of the 7th International Workshop on Artificial Intelligence and Statistics*.
- [141] Koo, J.-Y.; Sohn, I.; Kim, S.; and Lee, J. W. 2006. *Structured polychotomous machine diagnosis of multiple cancer types using gene expression*. *Bioinformatics* 22(8):950–958.
- [142] Yamada, M.; Sugiyama, M.; and Matsui, T. 2010. *Semisupervised speaker identification under covariate shift*. *Signal Processing* 90(8):2353–2361.
- [143] D. Collett, "Modelling binary data", London, Chapman & Hall, 369P, 1991.
- [144] J.C.V. Houwelingen, S. Le Cessie, "Logistic regression, a review". *Stat.Neerl.* 42,215-252., 1988.
- [145] D.R. Cox, "The analysis of binary data", London, Methuen, 1970.
- [146] D.R. Cox, E.J.Snell, "Analysis of binary data", London, Chapman & Hall, 236P, 1989.
- [147] F. Duyme et J.J. Claustrioux, "La régression logistique binaire", *Notes de statistique et d'informatique, Faculté universitaire des sciences agronomiques. Unité de statistique, informatique et mathématique appliquées, Gembloux (Belgique)*, 2006.
- [148] J.P. Nakache, J. Confais, *Statistique Explicative Appliquée, Partie 2, "Modèle Logistique"*. pages 77-168, Technip, 2003.

BIBLIOGRAPHIE.

- [149] R. Rakotomalala, "Pratique de la Régression Logistique: Régression Logistique Binaire et Polytomique". Version 2.0, Université Lumière Lyon 2, 27 Décembre 2009.
- [150] A. Rosenfeld, "From image analysis to computer vision: An annotated bibliography". 1955-1979. *Computer Vision and Image Understanding*, 84: 298_324, 2001.
- [151] V. Vapnik, "The Nature of Statistical Learning Theory". Springer, N.Y, 1995.
- [152] B. Scholkopf et A. J. Smola, "Learning with Kernels", MIT Press, 2002.
- [153] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, A.J. Smola and K.R. Mueller, "Constructing descriptive and discriminative non-linear features: Rayleigh coefficients in kernel feature spaces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004.
- [154] R. Héroult et Yves Grandvalet, "Régression Logistique Parcimonieuse". Manuscrit auteur, publié dans "9ème Conférence d'Apprentissage, Plate-forme AFIA", Grenoble, France, 2007.
- [155] G.C. Cawley et N.L.C. Talbot, "Sparse Bayesian Kernel Logistic Regression", *ESANN'2004 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium)*, 28-30 Avril 2004.
- [156] A.N. Tikhonov et V. Y. Arsenin, "Solutions of ill-posed problems". John Wiley, New York, 1977.
- [157] M.J. Jones and J.M. Rehg. "Statistical color models with application to skin detection". *International Journal of Computer Vision*, 46(1), pp. 81-96, January 2002.
- [158] C.A. Doukim, Jamal Ahmad Dargham, Ali Chekima et Sigeru Omatu, "Combining neural networks for skin detection". *Signal & Image Processing: An International Journal(SIPIJ)* Vol.1, No.2, December 2010.
- [159] Keerthi, S.; Duan, K.; Shevade, S.; and Poo, A. 2005. A fast dual algorithm for kernel logistic regression. *Machine Learning* 61(1-3):151–165.
- [160] P. Kakumanu, S. Makrogiannis, N. Bourbakis, "A survey of skin-color modeling and detection methods". *Pattern Recognition* 40- 1106 – 1122, 2007.

BIBLIOGRAPHIE.

- [161] J. Brand and J.S. Mason. "A comparative assessment of three approaches to pixel-level human skin-detection". In *Proceedings 15th International Conference on Pattern Recognition*, vol. 1, Barcelona, Spain, pp. 1056-1059, September 2000.
- [162] X. Fan, L. Wang, "Comparing linear discriminant function with logistic regression for the two-group classification problem", *J.Experim.Edu.*67, 265-286, 1999.
- [163] M. Abdullah-Al-Wadud, M. Shoyaib and O. Chae, "A Skin Detection Approach Based on Color Distance Map". *Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Article ID 814283*, 10 pages, 16 December 2008.
- [164] Z. Ma and A. Leijon, "Human skin color detection in RGB space with Bayesian estimation of beta mixture models". *Sound and Image Processing Lab KTH - Royal Institute of Technology, Aalborg, Denmark, August 23-27, 2010.*
- [165] D. Brown, I. Craw, and J. Lewthwaite, "A som based approach to skin detection with application in real time systems". In *Proc. of the British Machine Vision Conference*, volume 2, pp. 491-500, 2001.
- [166] B. Jedynak, H. Zheng, M. Daoudi and D. Barret, "Maximum entropy models for skin detection". *Tech. Rep. XIII, Universite des Sciences et Technologies de Lille, France, 2002.*
- [167] J.Y. Lee and S.I. Yoo, "An elliptical boundary model for skin color detection". In *Proc. International Conference on Imaging Science, Systems and Technology, Las Vegas, USA, June 2002.*
- [168] J. Brand and J.S. Mason. "A comparative assessment of three approaches to pixel-level human skin-detection". In *Proceedings 15th International Conference on Pattern Recognition*, vol. 1, Barcelona, Spain, pp. 1056-1059, September 2000.
- [169] Z. Ma and A. Leijon, "Human skin color detection in RGB space with Bayesian estimation of beta mixture models". *Sound and Image Processing Lab KTH - Royal Institute of Technology, Aalborg, Denmark, August 23-27, 2010.*
- [170] M.M. Aznaveh, H. Mirzaei, E. Roshan, and M. Saraee, "A new and improved skin detection method using RGB vector space". In *Proceedings of IEEE International Multi- Conference on Systems, Signals and Devices, July 2008*, pp.1-5.

BIBLIOGRAPHIE.

- [171] T.Y. Chin, "Fuzzy Skin detection". A project for the award of the degree of Master of Science (Computer Science) in Faculty of Computer Science and Information System Universiti Teknologi Malaysia, October 2008.
- [172] SHARMA, G. et TRUSSELL, H. (1942). Digital color imaging. *IEEE Transactions on Image Processing*, 32:247–274.
- [173] BERRY, D. (1987). Colour recognition using spectral signatures. *Pattern Recognition Letters*, 6:69–75.
- [174] BRUN, L. et TREMEAU, A. (2002). Color quantization, chapitre 9, pages 589–637. *Electrical and Applied Signal Processing Series*. CRC Press.
- [175] TRÉMEAU, A., FERNANDEZ-MALOIGNE, C. et BONTON, P. (2004). *Image numérique couleur : de l'acquisition au traitement*. Dunod.
- [176] G. Rougeron. *Problèmes liés à la couleur en synthèse d'images*. Thèse de doctorat, Université Jean Monnet de Saint-Etienne, janvier 1998.
- [177] C. C. Brunner, A. G. Maristany, D. A. Butler, D. Vanleuween, et J. W. Funck. An evaluation of color spaces for detecting defects in Douglas-fir veneer. *Industrial Metrology*, 2(3 and 4) :169–184, 1992.

BIBLIOGRAPHIE.

BIBLIOGRAPHIE.

BIBLIOGRAPHIE.

BIBLIOGRAPHIE.

BIBLIOGRAPHIE.

BIBLIOGRAPHIE.
