



HAL
open science

Segmentation du masque capillaire dans un visage

Cedric Rousset

► **To cite this version:**

Cedric Rousset. Segmentation du masque capillaire dans un visage. Interface homme-machine [cs.HC]. Institut National Polytechnique de Grenoble - INPG, 2010. Français. NNT: . tel-00530635

HAL Id: tel-00530635

<https://theses.hal.science/tel-00530635>

Submitted on 29 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE DE GRENOBLE
INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

THÈSE

pour obtenir le grade de

DOCTEUR DE L'Université de Grenoble
délivré par l'Institut Polytechnique de Grenoble

Spécialité : "Signal, Image, Parole, Télécommunications"

préparée au laboratoire Grenoble Image Parole Signal et Automatique

dans le cadre de l'école **Doctorale EEATS**
"Électronique, Électrotechnique, Automatique, Télécommunications, Signal"

présentée et soutenue publiquement

par

Cédric ROUSSET

le 12 / 10 / 2010

Titre :

SEGMENTATION DU MASQUE CAPILLAIRE DANS UN VISAGE

Directeur de thèse : Pierre-Yves COULON

JURY

Madame	Michèle ROMBAUT,	Présidente
Monsieur	Alain TREMEAU,	Rapporteur
Monsieur	Christian GERMAIN,	Rapporteur
Monsieur	Pierre-Yves COULON,	Directeur de thèse
Monsieur	Christophe BLANC,	Examineur

Remerciements

Les travaux de thèse présentés dans ce manuscrit ont été effectués au Département Images et Signal (DIS) du laboratoire GIPSA-Lab. C'est avec plaisir que j'exprime ici mes remerciements à toutes les personnes qui ont contribué, de près ou de loin à ce document durant ces trois années.

Tout d'abord, je tiens à remercier les membres du jury d'avoir été présents lors de la soutenance malgré les perturbations dans les transports occasionnées par les grèves. Je remercie Monsieur Germain et Monsieur Tremeau, qui m'ont fait l'honneur d'accepter de relire ce présent document et d'apporter respectivement une certaine texture et une couleur à ce manuscrit par leurs observations et remarques pertinentes. Je remercie également Monsieur Blanc pour ne pas avoir maquiller son intérêt qu'il a porté à ce manuscrit en l'examinant et en assistant à la soutenance. Mes remerciements vont aussi à Michèle Rombaut pour avoir présidé la soutenance et d'avoir eu de la croyance transférable pour mes travaux et qui a contribué à ce manuscrit par ses connaissances en fusion de données.

Je tiens tout particulièrement à remercier mon directeur de thèse, Pierre-Yves Coulon, pour ses conseils avisés qui m'ont plus d'une fois permis de "trouver" des nouvelles pistes de recherche et sa patience pour avoir écouté parfois mes théories les plus capillo-tractées. Je le remercie aussi pour sa confiance durant ces trois ans ainsi que pour sa grande disponibilité.

Je souhaite remercier aussi tous les membres du GIPSA-lab, permanents et thésards qui ont rythmé mes années de recherche sur les sites de Viallet et du Campus. Que ce soit scientifiquement ou humainement parlant, j'ai eu plaisir à partager ces nombreux moments d'échanges. Je remercie notamment les doctorants présents et passés, les coincheurs de midi et la fine équipe de la Roumanie : Nico, Julien et les autres. Je remercie, tout particulièrement, mes cobureaux qui ont dû me supporter plus que les autres et réciproquement : Barthélémy et son téléphone portable, Bertrand pour son "*énergie*" et son amour pour David Beckham et Christian, le métalleux qui aura gardé ses cheveux jusqu'au bout. Je remercie aussi tous ceux qui ont participé pendant ces trois ans à mon équilibre psychologique quand je m'arrachai les cheveux sur ma thèse : L'équipe de la BDM : Thomas, Benoit, Melissa ..., La troupe du O Calaghan : Pauline, Cha, Nath, Ben, François et la "*Hot bar tender*", les anciens de L'Isen, Les Alpha légionnaires : Kheb le nabab du khebab et Carine, Pierrot, Auré l'ours de la tarentaise, Le malin Manu, DamDam, Le Kiross 25, Shagguy , Djé, Narrach Warpaper ..., La team du PSHT : Boiss le Kiwi fou de la fléchette, Bob le canonnier, Batto le magicien du Texto, KB, JP, Gaby, Francho, Pika, Fab. Je remercie aussi tous ceux que j'aurais oublié de citer mais qui ont été là.

Je remercie enfin les membres de ma famille qui ont toujours été avec moi et particulièrement mes parents qui m'ont toujours soutenu et poussé pour aller toujours plus loin.

Enfin je te remercie toi, lecteur, qui a pris la peine de lire ces remerciements jusqu'au bout et qui je l'espère lira un peu plus de ce manuscrit.

Résumé

Cette thèse porte sur la segmentation du masque capillaire dans un visage. Nous présentons dans ce manuscrit une étude sur la localisation et la caractérisation d'une chevelure dans une image. L'objectif de ces travaux est de proposer une méthode permettant de définir automatiquement la zone capillaire avec la meilleure fiabilité possible afin de prendre en compte la grande variabilité de la représentation d'une chevelure.

La segmentation est effectuée par un traitement "*Matting*" qui est une méthode de segmentation par approche "*régions*". Cet algorithme sépare une image en deux plans : un plan d'information qui représente la chevelure et un plan de fond. Il est initialisé par la définition de marqueurs qui vont être diffusés dans l'ensemble de l'image. L'obtention d'une bonne segmentation dépend directement de la précision dans le placement de ces marqueurs. Nous définissons leurs positions par une analyse conjointe de trois paramètres caractéristiques d'une chevelure : sa texture, sa teinte et sa position.

Dans un premier temps, nous avons mis en place une analyse fréquentielle pour caractériser la texture. Nous utilisons un filtrage de l'image par un filtre passe-bande gaussien isotrope. Nous modélisons un masque de localisation des zones fréquentielles similaires à la chevelure. Dans un second temps, nous avons mis en place une analyse couleur pour caractériser la teinte. Nous utilisons un classifieur couleur qui représente la distribution du modèle couleur par une distribution gaussienne sur chaque composante de chrominance dans l'espace couleur YCbCr.

Une première information commune de localisation d'une chevelure est définie et est combinée par une méthode de fusion de données basée sur le modèle des fonctions de croyances transférables. Cette approche permet notamment la prise en compte de différents degrés d'ignorance par la modélisation d'un état d'"*incertitude*". L'ajout de ce nouvel état est particulièrement adapté à notre algorithme de segmentation puisqu'il permet de contrôler la position des pixels dont l'état est estimé pendant le traitement "*Matting*". Cette approche est ensuite optimisée par l'ajout d'une fonction d'affaiblissement basée sur la caractérisation de la localisation d'une chevelure autour du visage qui est le troisième paramètre capillaire caractéristique. Cette fonction permet de pondérer la fiabilité de nos sources de données par rapport à l'éloignement du visage. En effet plus un pixel est loin d'un visage moins il a de probabilité d'appartenir au masque capillaire.

La segmentation capillaire est évaluée grâce à une analyse quantitative de la définition d'un masque de chevelure en comparaison avec une base de référence obtenue par une segmentation semi-manuelle. Enfin le masque capillaire est caractérisé par trois descripteurs, un pour chaque paramètre capillaire. Cette classification permet la description d'une chevelure par une approche similaire à une description cognitive faite par un observateur humain.

Abstract

This thesis focuses on segmentation of the hair mask in a face. This report presents the study on the localization and the characterization of hair representation in a picture in a picture. The aim of the thesis is to propose a automatic method to segment the hair with best possible reliability in order to take into account the variability of different hairstyles.

The segmentation is performed by a *Matting* treatment which is a region-contour based approach. This algorithm extract a foreground object from an image. The picture is divided into 2 plans : the first one represents the area of hair information and the second one represents the background plan. This approach is initialized by the definition of some markers which will be diffused into the entire picture. Obtaining a good segmentation depends directly on the precision of these markers placement. We define their position by a combined analysis of three parameters which are characteristics of hair : its texture, its color and its position around the face.

In a first time, we set up a frequential analysis to characterize the hair texture. We performed a filtering of the image using a Gaussian band pass isotropic filter. We define a localization mask based on the detection of the frequential zones which are similar to hair. In a second time, we set up a hair color analysis. We define a color classifier which represents the distribution of the color model by a Gaussian distribution on each component of chrominance in YCbCr color space.

Common information of hair localization are combined by a method of data fusion based on the transferable belief model. This approach allows to take into account various degrees of ignorance by the modelization of a state of "*uncertainty*". The addition of this new state is well adapted to our algorithm of segmentation since it makes possible to control the position of the pixels whose state is estimated during the *Matting* treatment. Then this approach is improved by the addition of a discounting function based on the third hair parameter which is the localization of hair around the face. This function leads to balance the reliability of our sources compared to the distance of the face. Indeed the probability for a pixel to belongs to the hair mask decreases as the pixel is far from the face.

The hair segmentation is evaluated thanks to a quantitative analysis by a comparison with a baseline of references hair masks obtained by a semi-manual segmentation. Finally the hair mask is characterized by three descriptors for each hair parameter. This classification allows hair description by an approach similar to a cognitive description made by a human observer.

Table des matières

1	Les Reflets de la Chevelure	3
1.1	La coiffure, reflet de la personnalité	5
1.2	La coiffure, reflet d'identité	9
1.3	La coiffure, reflet du réalisme	14
1.4	Conclusion	18
2	La Segmentation Capillaire	19
2.1	État de l'Art sur l'Analyse Capillaire	21
2.2	Techniques de Segmentation	34
2.3	Conclusion	39
3	Cadre de la Segmentation Capillaire	41
3.1	Digital Matting ou Composition d'images	43
3.2	"Matting" par marqueur	49
3.3	Etude du "Matting" par marqueur	52
3.4	Cadre d'étude	57
3.5	Évaluation des résultats	62
3.6	Conclusion	64
4	Analyse Fréquentielle	65
4.1	Analyse de texture	67
4.2	Analyse Fréquentielle de la Chevelure	74
4.3	Définition automatique des paramètres	84
4.4	Conclusion	92
5	Analyse Couleur	95
5.1	Analyse Couleur	97
5.2	La Segmentation Couleur supervisée	101
5.3	Segmentation couleur de la peau	106
5.4	Segmentation couleur de la chevelure	115
5.5	Conclusion	116
6	Fusion de données	119
6.1	La Fusion de Données	121
6.2	Les Fonctions de croyances	126
6.3	La Fusion Capillaire	132
6.4	Conclusion	150

7	Caractérisation du masque capillaire	153
7.1	Caractérisation de la Texture	155
7.2	Caractérisation de la Teinte	158
7.3	Coloration du masque capillaire	160
7.4	Caractérisation de la Forme	163
7.5	Conclusion	169
A	Le Cheveu	179
A.1	Le Cheveu	181
A.2	Structure	182
A.3	Cycle de vie	185
A.4	Type de Chevelure	186
A.5	Coloration	187
A.6	Propriétés du Cheveu	190
B	La Rétine	193
B.1	Les photo-récepteurs	196
B.2	La Couche Plexiforme Externe	196
B.3	La Couche Plexiforme Interne	197
B.4	Cortex <i>V1</i>	198
C	Les Espaces Couleurs	201
C.1	L'Espace <i>RVB</i>	203
C.2	Les Espaces de Télécommunication : <i>YCbCr</i> - <i>YUV</i> - <i>YIQ</i>	204
C.3	Les Espaces Perceptuels <i>HSV</i> - <i>HSL</i>	205
C.4	Les Espaces Colorimétriques <i>XYZ</i> - <i>La*b*</i>	207

Introduction

La chevelure a une place importante dans notre société. Elle est un élément important de l'image d'une personne. C'est un reflet de sa personnalité qui lui permet d'exprimer son appartenance sociale, culturelle ou idéologique. Elle intervient aussi dans le processus de reconnaissance faciale d'une personne où elle apporte une part importante d'information sur l'identité. Elle est aussi un élément distinctif lors de la description d'une personne : souvent la couleur des cheveux et le type de coupe sont détaillés. Enfin elle a une grande place dans la mode et fait l'attention de nombreux soins avec la profusion de cosmétiques pour son entretien.

Problématique

Paradoxalement, contrairement à d'autres éléments du visage, elle est très peu étudiée et utilisée pour des applications numériques. Ce n'est que ces dernières années que l'intérêt pour la chevelure est apparu, au départ, principalement pour des applications de synthèses d'images et depuis peu pour des applications de réalité augmentée ou des applications biométriques. La principale limite, à l'origine du faible nombre de travaux traitant de la chevelure, est sa grande variabilité. Contrairement à d'autres caractéristiques faciales (yeux, nez, bouche), une chevelure peut prendre de nombreuses teintes, formes et positions. Cette grande variabilité va devoir être prise en considération pour le développement d'une application de segmentation et de caractérisation. Nous allons devoir mettre au point un traitement qui pourra s'adapter au plus grand nombre de configurations capillaires possibles.

Objectif

L'objectif de cette thèse est de proposer une technique de segmentation automatique du masque capillaire. Comme nous venons de le dire, la grande difficulté de cette étude est de proposer une méthode qui permet de traiter le maximum de cas possibles malgré la grande variabilité de la chevelure. C'est pourquoi il faut déterminer les paramètres capillaires les plus stables pour l'ensemble des coupes qui permettent la description et la localisation d'une chevelure dans une image. Il faut aussi définir une méthode de fusion qui permet de combiner au mieux l'information obtenue à partir de ces différents paramètres. Enfin, il faut mettre en place un algorithme de segmentation performant et robuste c'est à-dire qui permette la définition fidèle d'un masque de position de la chevelure. Ces travaux de thèse s'inscrivent dans une démarche innovante car nous sommes encore dans les prémices de la recherche sur la segmentation capillaire. Mais c'est aussi ce qui en fait sa difficulté puisque aucun cadre de travail formel n'est établi. Ainsi de nombreuses pistes de recherches restent encore à explorer.

Contribution

Dans cette thèse, nous avons modélisé une approche de segmentation automatique de la chevelure par croissance de région. Cette segmentation est basée sur l'étude de 3 paramètres caractéristiques de la chevelure : sa texture, sa couleur et sa position autour du visage. Les principales contributions de cette thèse sont :

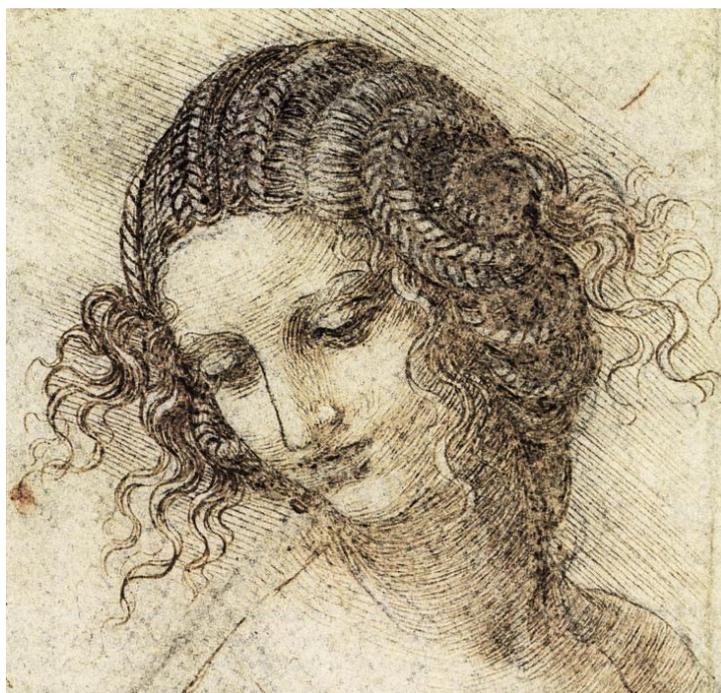
- L'utilisation d'un troisième paramètre pour localiser la chevelure : sa texture, avec la mise en place d'un filtrage pour segmenter le maximum d'information fréquentielle similaire à la chevelure.
- L'utilisation des fonctions de croyances pour le processus de fusion de données
- L'utilisation de l'algorithme de "*Matting*" pour la segmentation de l'image.
- La mise en place d'une nouvelle méthode de création de bases de référence. Méthode semi-manuelle qui combine une segmentation automatique initialisée par l'apport manuel de connaissance par un expert.
- La définition de descripteurs permettant de caractériser un masque capillaire selon ses 3 paramètres caractéristiques.

Organisation du manuscrit

Ce manuscrit présente les travaux réalisés dans le cadre de ma thèse effectuée au sein du laboratoire GIPSA-Lab à Grenoble. Il est constitué de 7 chapitres :

- Le premier chapitre présentera le contexte de recherche et l'intérêt de l'analyse capillaire aux travers des diverses applications où l'on pourrait l'utiliser : applications de réalité augmentée, applications biométriques et applications de synthèse d'images.
- Le chapitre 2 sera consacré, tout d'abord, à l'état de l'art sur l'analyse et la segmentation capillaire, suivi d'une étude sur les techniques de segmentation par approche "contours" et par approche "régions".
- Dans le chapitre 3, nous présenterons plus précisément l'algorithme de segmentation par croissance de régions : le traitement "*Matting*". Nous ferons une étude sur ses caractéristiques d'initialisation et les limites de son application. Cette étude nous permettra de définir un cadre d'étude pour notre segmentation capillaire.
- Dans le chapitre 4, nous présenterons l'étude pour caractériser la texture de la chevelure. Nous présenterons le filtre fréquentiel mis en place et l'étude pour le paramétrer automatiquement.
- Dans le chapitre 5, nous présenterons notre étude couleur sur la teinte de la chevelure. Nous présenterons le choix de l'espace et du classifieur utilisés et leurs applications pour la segmentation couleur du masque de peau du visage et du masque capillaire.
- Dans le chapitre 6, nous présenterons notre approche de fusion des données. Dans un premier temps, nous verrons comment réunir l'information fréquentielle et couleur par le modèle des fonctions de croyances. Dans un second temps, nous verrons comment nous avons intégré l'information de la localisation de la chevelure par une fonction d'affaiblissement.
- Finalement dans le chapitre 7, nous présenterons les premières pistes de recherches pour la caractérisation du masque capillaire.

1 Les Reflets de la Chevelure



Sommaire

1.1	La coiffure, reflet de la personnalité	5
1.1.1	Le Maquillage Virtuel	7
1.1.2	Le projet Makeuponline	8
1.2	La coiffure, reflet d'identité	9
1.2.1	Identification & authentification de personne	13
1.3	La coiffure, reflet du réalisme	14
1.3.1	La Synthèse de visage	15
1.3.2	La Synthèse Capillaire	17
1.4	Conclusion	18

"La coiffure est l'art d'arranger les cheveux, éventuellement de modifier leur aspect extérieur".

Cet art implique différents types d'interventions : couper les cheveux, les friser ou les défriser, les tresser, les teindre, les décolorer, les poudrer, les enduire de substances variées, leur adjoindre de faux cheveux (sous la forme d'une perruque ou d'un postiche ou rallonges capillaires qui consiste à garder les vrais cheveux et à seulement les allonger) ou les orner de façons plus ou moins sophistiquées selon les cultures et les civilisations¹.

1.1 La coiffure, reflet de la personnalité

Plus qu'un simple détail de l'apparence humaine, la chevelure est en réalité un élément essentiel de la personne. Célébrés de tous temps, les cheveux sont le reflet d'une mode, d'une tradition ou d'une personnalité.



(a) La Dame de Brassempouy

(b) Tresse Egyptienne

(c) Coiffure Himba

FIGURE 1.1 – **La Coiffure, reflet de la personnalité 01** : (a) Sculpture de "La Dame de Brassempouy", (b) Enfant Egyptien arborant la tresse typique des jeunes égyptiens au temps des pharaons, (c) Femme Himba arborant sa coiffure teintée symbole de fertilité.

Dès la préhistoire, les femmes entretenaient leurs chevelures en recourant aux tresses. De nombreuses statuettes trouvées sont des représentations de femmes portant des coiffures. La plus ancienne statuette connue est la sculpture de "La Dame de Brassempouy". Fragment en ivoire, datant du Paléolithique supérieur, elle représente un visage humain coiffé d'un maillage de lignes croisées. Longtemps considérée comme une capuche, il semblerait que ce soit plus vraisemblablement la représentation d'une coiffure tressée (*figure 1.1a*).

Au cours du temps le rôle de la chevelure a évolué. Autrefois, la coiffure au même titre que les vêtements, pouvait représenter l'appartenance à certaines classes sociales ou l'évolution par certains stades de la vie (jeunesse, puberté, âge adulte). Ainsi par exemple, au temps des pharaons, les jeunes égyptiens étaient tous coiffés d'une unique tresse jusqu'à leur passage à l'âge adulte où les garçons devaient se raser et les filles se laissaient pousser leurs cheveux (*figure 1.1b*).

La chevelure est souvent associée à la fertilité, car elle pousse et se régénère sans cesse. Dans de nombreuses tribus africaines, elle est comparée à un buisson qu'il faut entretenir et orner. La

1. source : Wikipédia

coiffure est l'objet de grands soins et est souvent colorée grâce à des pigments naturels. Encore aujourd'hui certaines tribus ont conservé cette symbolique, ainsi les femmes des tribus Himba en Namibie se teignent le corps et les cheveux en rouge avec une pommade réalisée à base de graisse animale et de poudre d'hématite (*figure 1.1c*). De nombreuses autres traditions et significations peuvent être attribuées à la chevelure que le lecteur intéressé pourra trouver dans l'ouvrage "Les vies du cheveu" [Auzou 01].



(a) Louise Brooks



(b) Groupe de Punk



(c) Membre du groupe Iron Maiden

FIGURE 1.2 – La Coiffure, reflet de la personnalité 02 : (a) Louise Brooks symbole du féminisme, (b) Groupe de punk arborant des coiffures en forme de crête, (c) Membres du groupe de Métal Iron Maiden arborant de longues chevelures.

De nos jours, la chevelure sert plus à exprimer sa différence ou son originalité. La coiffure ne reflète plus une appartenance publique mais une expression plus individuelle de sa personnalité. Même s'il est toujours soumis à certains codes d'appartenances, le choix d'une chevelure n'est plus imposé par des règles sociales mais bel et bien par un choix individuel. On peut notamment citer Louise Brooks qui en 1929 se pare d'une coupe courte dite "à la garçonne" qui représente une libération de la femme et un symbole du féminisme (*figure 1.2a*). On peut aussi citer les coiffures excentriques du mouvement Punk qui exprime une révolte contre les valeurs établies (*figure 1.2b*) ou les longues coiffures arborées par les fans de musique Métal (*figure 1.2c*).

La couleur des cheveux, leur forme, ou encore le style de coupe adopté sont autant d'éléments importants de notre physique qui influent dans la vie de tous les jours. Sans forcément être excentrique, notre chevelure de tous les jours joue un rôle sur la perception qu'ont les autres

de nous-mêmes et de notre personnalité. La chevelure fait toujours l'objet de tous les soins, encore plus dans notre société moderne avec le développement de la cosmétique. Récemment de nombreux studios de maquillage virtuel ont d'ailleurs fait leur apparition sur Internet afin de pouvoir proposer aux acheteurs potentiels la possibilité de tester les gammes de produits de maquillage toujours de plus en plus importantes.

1.1.1 Le Maquillage Virtuel

Le maquillage virtuel est une des applications qui a subi un grand essor avec l'avènement de l'internet [LOreal] [Gemey] [Makeover] [BeauteFamille] [Taa] [Vesalis]. À partir d'une photo d'un modèle ou d'une photo de soi téléchargée, les utilisateurs peuvent appliquer différents maquillages afin d'essayer une large palette de cosmétiques et ainsi avoir une idée de l'apparence qu'il pourrait avoir. Les techniques de biométrie ont permis de rendre en partie automatique la détection des paramètres faciaux par exemple la détection des lèvres pour une application de rouge à lèvres/gloss ou la détection des yeux pour des applications de fard ou de mascara ...(*figure 1.3*)

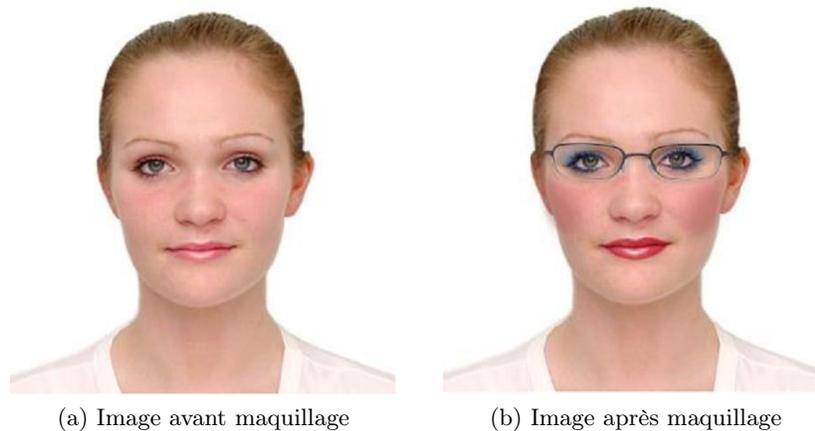


FIGURE 1.3 – **Maquillage virtuel - Application** : Exemple de maquillage virtuel avec ajout de mascara, fard à paupières et fard à joue, de rouge à lèvres ainsi que d'accessoires comme les lunettes [BeauteFamille]

Une des applications présente est notamment la possibilité de changer la couleur ou de changer la coupe de la chevelure et d'essayer différents modèles de coiffures classiques (*figure 1.4a*), mais aussi des modèles de coiffures tirées de photo de stars (*figure 1.4b*). La technique généralement utilisée est l'ajout d'une perruque virtuelle directement plaquée par-dessus le visage à maquiller. Si ces coiffures virtuelles donnent des résultats plutôt convaincants, elles ne sont pas totalement adaptées aux visages, les décalages de taille et de position de la chevelure par rapport au visage perturbent la bonne perception et reconnaissance du visage (*figure 1.4c*). Il existe toujours la possibilité de recadrer le masque capillaire afin de l'adapter au visage : adaptation en taille (zoom), adaptation en longueur ou en largeur et adaptation en rotation, mais ces ajustements ne prennent généralement pas en compte l'interaction entre la chevelure et le visage, le cou ou les épaules. De plus, les mannequins modèles ont toujours les cheveux attachés et tirés en arrière ce qui laisse une faible partie de chevelure à recouvrir et qui permet d'éviter de devoir retirer la chevelure déjà présente.

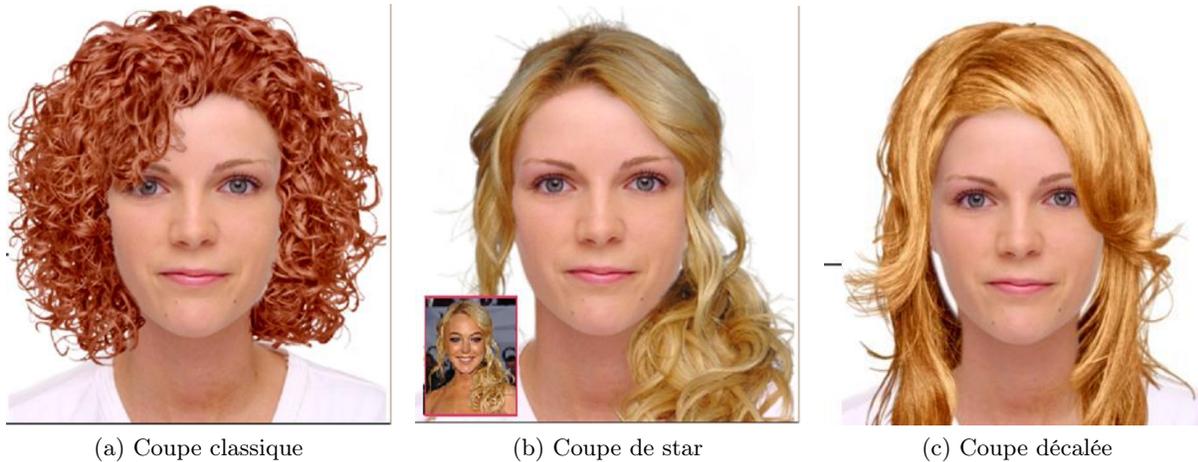


FIGURE 1.4 – **Maquillage virtuel - Changement de Chevelure** : Exemples d'applications (a) Ajout d'un modèle générique de coiffure, (b) Ajout d'un modèle de coiffure de star (ici Lindsay Lohan), (c) Ajout d'un modèle générique de coiffure mal adapté. [BeauteFamille]

Il existe aussi la possibilité de changer la teinte de sa chevelure directement sans devoir rajouter de perruque virtuelle (*figure 1.5*). La chevelure doit alors être segmentée grossièrement manuellement par l'utilisateur. Le changement de teinte possible de la chevelure résulte de l'application de modèle de couleur prédéfinie directement sans prise en compte de la couleur initiale. Il en résulte une coloration qui n'est pas naturelle sans prendre en compte les zones de reflet ou les zones d'ombres de la chevelure pour adapter le modèle couleur à ces zones.

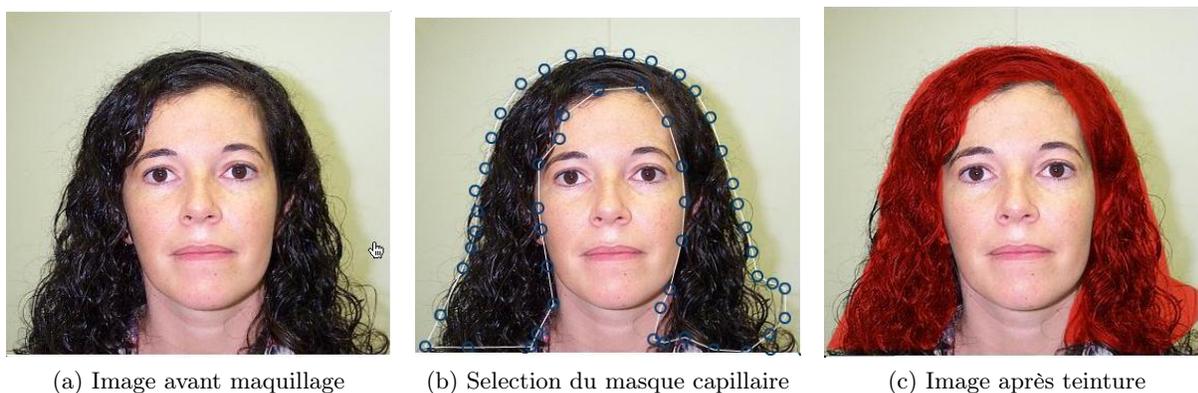


FIGURE 1.5 – **Maquillage virtuel - Coloration de Chevelure** : Exemple d'application (a) Image avant maquillage, (b) Image après sélection semi-manuelle de la position de la coiffure, (c) Image après application du changement de teinte [Coiffeursdefrance].

1.1.2 Le projet Makeuponline

La société Vesalis [Vesalis] est une start-up française qui travaille en partenariat avec le GIPSA-lab. Cette entreprise est spécialisée dans le développement d'applications pour la réalité

augmentée. Elle développe actuellement un simulateur virtuel de mise en beauté automatique, temps réel et interactif (*figure 1.6*). Contrairement aux autres applications de maquillage virtuel, Vesalis travaille sur une application entièrement automatisée où l'action de l'utilisateur doit être limitée aux choix de maquillage à tester.

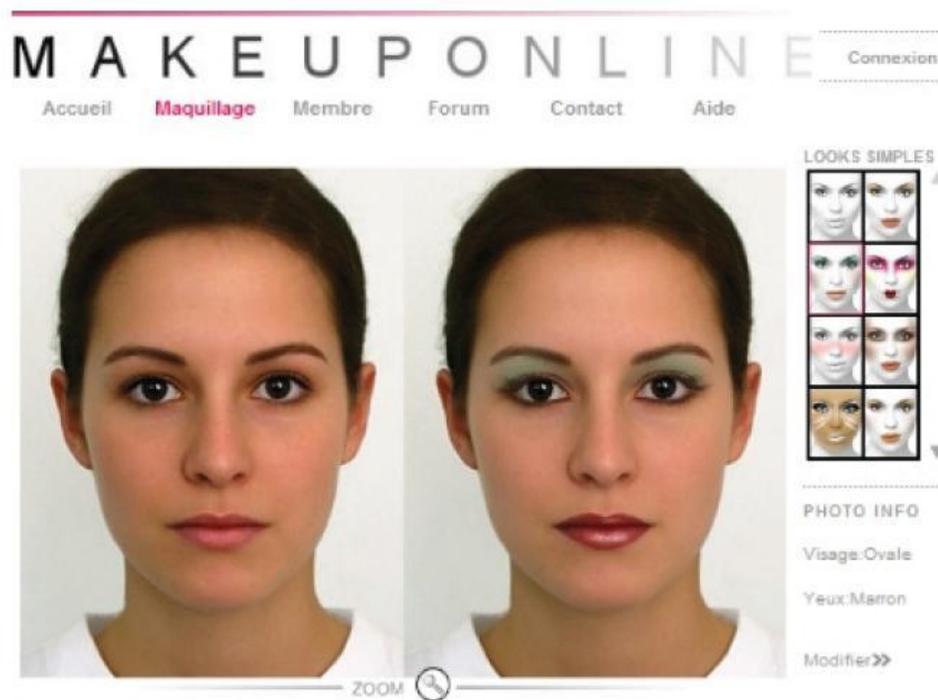


FIGURE 1.6 – **Projet Makeup Online de la Société Vesalis** : Interface de la plate forme de maquillage virtuel *Makeup online* de la société Vesalis [Vesalis].

La prise en charge de la chevelure est une piste de développement de leur application. L'analyse et la segmentation du masque capillaire va permettre dans un premier temps une segmentation automatique de la chevelure sans que l'utilisateur ait besoin d'intervenir. Cette tâche automatisée va permettre une segmentation plus rapide et plus fine du masque capillaire. L'analyse du masque capillaire segmenté permettra soit une étude colorimétrique de la chevelure pour améliorer l'application d'une teinture, soit une étude morphologique de la chevelure pour déterminer des caractéristiques de la chevelure par rapport au visage afin d'adapter le plus naturellement possible les perruques virtuelles au visage.

L'étude et l'analyse de la chevelure ouvrent aussi des débouchés pour de nombreuses applications autres que les applications de réalité augmentée. Entre autre le développement de la segmentation capillaire peut aussi apporter des pistes d'améliorations pour des applications biométriques de reconnaissances faciales.

1.2 La coiffure, reflet d'identité

La chevelure représente une composante de l'apparence humaine. La reconnaissance d'une personne n'est pas seulement basée sur l'analyse de son visage mais la chevelure intervient également dans ce processus. Par exemple, nous sommes capables de reconnaître une personne connue

se trouvant de dos grâce à plusieurs éléments : sa stature, sa façon de bouger et sa chevelure. Même quand les éléments du visage sont visibles, la chevelure apporte encore une information importante sur l'identité d'une personne.

Le processus cognitif par lequel le cerveau analyse une image pour y détecter et identifier un visage s'appelle "La perception des visages" (*Face perception*). L'être humain et certains primates ont développé une faculté à identifier très rapidement un visage dans son environnement. De plus l'être humain est capable de reconnaître l'identité d'une personne parmi plusieurs centaines d'autres. Cette aptitude repose sur des mécanismes neurocognitifs, en partie innés, dont certains sont présents chez le nourrisson dès sa naissance². Ce caractère spécifique du processus de reconnaissance faciale a été mis en évidence par plusieurs pathologies humaines qui sont à l'origine d'anomalies dans la perception des visages. Elles peuvent être d'origine neurophysiologique, traumatique ou psychiatrique.

La plus connue est la *prosopagnosie* qui prive ceux qui en souffrent de la capacité à reconnaître les visages et uniquement les visages. Ces personnes ne discernent plus un individu d'un autre par ses traits faciaux mais seulement par le son de sa voix ou par un détail particulier qui le caractérise (une tâche de naissance, une paire de lunettes,...). Même si les personnes souffrant de *prosopagnosie* arrivent parfaitement à évaluer l'âge, le sexe et l'ethnie d'un visage, il semble qu'elles perdent la notion des mensurations anthropométriques du visage telles que la taille du nez, l'épaisseur des lèvres, la forme des yeux dans le processus de reconnaissance faciale, cependant ils peuvent reconnaître des changements d'expression ou de direction du regard. Ce qui est à noter c'est que cette perte de distinction ne s'applique qu'aux visages. En effet plusieurs études sur des personnes souffrant de *prosopagnosie* ont prouvé que ces personnes étaient capables d'identifier et de reconnaître des formes complexes. Un exemple mettant en évidence cette particularité est le cas d'un berger qui malgré sa *prosopagnosie* était capable de reconnaître la plupart des moutons de son troupeau [McNeil 94]. Cette pathologie montre bien la spécialisation de l'homme dans la tâche de reconnaissance de personne et la modélisation numérique d'un tel processus n'est pas simple.

Dans [Jarudi 07], Jaruddi *et al* ont étudié l'apport des éléments internes du visage (yeux, nez, bouche...) par rapport aux éléments externes (Chevelure, forme du visage...). Ils ont particulièrement mis en évidence l'apport des éléments externes du visage par 2 expériences. La première évalue la part du rôle de chaque partie du visage dans le processus de reconnaissance pour différents niveaux de dégradation de l'image en la rendant plus ou moins floue. Ils ont défini 4 catégories (*figure 1.7*) : les éléments internes du visage placés en ligne (1), les éléments internes du visage placés selon leurs configurations initiales dans le visage (2), les éléments externes du visage auxquels ont été enlevés les éléments internes (3), le visage complet comprenant les éléments internes et externes (4).

L'expérience consiste à la reconnaissance de personnes célèbres pour chacune des 4 catégories avec des images plus ou moins floue (*figure 1.8*). Les résultats montrent que les visages complets (4) donnent les meilleurs résultats de reconnaissance faciale quel que soit la dégradation de l'image. Pour les autres catégories sans dégradation, les éléments internes du visage sont les plus efficient avec environ 80% de bonnes reconnaissance loin devant la chevelure qui obtient seulement 40%. Néanmoins, on constate qu'après 3 dégradations par floue les résultats de détec-

2. source Wikipédia

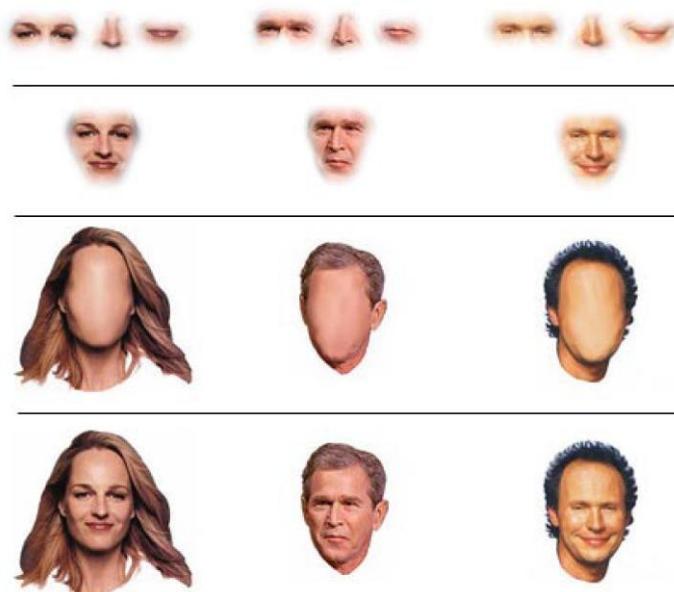


FIGURE 1.7 – **Importance de la chevelure dans le processus de reconnaissance faciale - Protocole expérience 1** [Jarudi 07] : Exemples des 4 catégories : 1^{ère} ligne, les éléments internes du visage placés en ligne (1), 2^{ème} ligne, les éléments internes du visage placés selon leurs configurations dans un visage (2), 3^{ème} ligne, les éléments externes du visage auxquels on a enlevé les éléments internes (3), 4^{ème} ligne, le visage complet comprenant les éléments internes et externes (4).

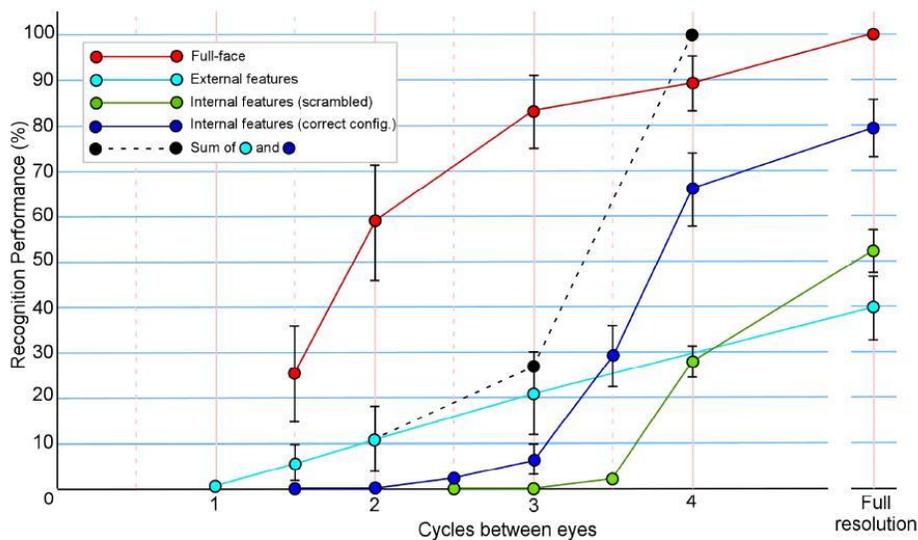


FIGURE 1.8 – **Importance de la chevelure dans le processus de reconnaissance faciale - Résultats expérience 1** [Jarudi 07] : Résultats pour les 4 catégories en fonction du nombre de dégradation de l'image : en vert, les éléments internes du visage placés en ligne (1), en bleu, les éléments internes du visage placés selon leurs configurations dans le visage (2), en cyan les éléments externes du visage auxquels on a enlevé les éléments internes (3) et en rouge, le visage complet comprenant les éléments internes et externes (4).

tion par les éléments du visage que ce soit en ligne (1) ou correctement disposé (2) s'effondrent alors que la reconnaissance par chevelure décroît doucement. Ces résultats montrent que la reconnaissance faciale se base plus sur les éléments externes pour des conditions de vue dégradée. Ces résultats montrent aussi que la fusion de données effectuée entre les informations apportées par les éléments externes et internes du visage n'est pas linéaire, résultat d'autant plus mis en évidence pour un fort niveau de dégradation de l'image où par exemple les éléments internes et externe combinés permettent environ 60% de reconnaissance alors qu'ils permettent une reconnaissance quasi nulle séparément.

Ainsi même si individuellement chaque élément du visage apporte très peu d'information sur l'identité d'une personne, il reste néanmoins nécessaire comme référence pour la configuration spatiale du visage. Cette analyse est mise en évidence par la deuxième expérience effectuée (*figure 1.9*) où le rapport entre les éléments internes et externes est évalué. Dans cette expérience, la résolution de la taille des éléments internes et externes du visage est modifiée tour à tour. Nous obtenons à partir d'un visage, 3 nouvelles configurations possibles : un visage où seuls les éléments internes ont été modifiés, un visage où seuls les éléments externes ont été modifiés et un visage où à la fois les éléments internes et externes ont été changés (*figure 1.9*). On constate que les modifications appliquées à la fois aux éléments externes et internes perturbent beaucoup moins la reconnaissance d'un visage que la seule transformation de taille d'un des 2 éléments. Ainsi la reconnaissance d'un visage résulte de l'interaction entre les éléments internes et externes du visage.

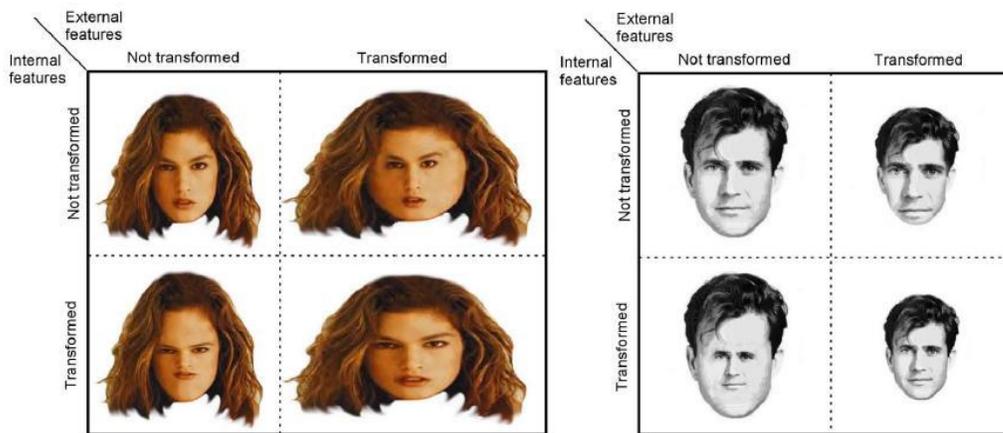


FIGURE 1.9 – **Importance de la chevelure dans le processus de reconnaissance faciale - Expérience 2** [Jarudi 07] : En haut à gauche le visage de référence, en haut à droite déformation des éléments externes du visage, en bas à gauche déformation des éléments internes du visage, en bas à droite déformation identiques des éléments internes et externes du visage.

L'importance des éléments externes du visage a été aussi mise en évidence par Sinha *et al* [Sinha 96], grâce à leur illusion d'optique appelée "*Illusion présidentielle*" (*Fig 1.10*). Cette illusion représente Bill Clinton et Al Gore, 2 hommes politiques américains, vue de face. On différencie parfaitement les 2 hommes bien que la reconnaissance et la différenciation ne se basent uniquement que sur les éléments externes de leurs visages : la forme du visage et particulièrement la coupe de cheveux. En effet, les éléments internes du visage de Bill Clinton ont remplacés ceux du visage de Al Gore.



FIGURE 1.10 – **Illusion présidentielle** : Illusion d'optique illustrant l'importance de la chevelure dans le processus de reconnaissance faciale. Les éléments du visage de Bill Clinton ont été intégrés au visage de Al Gore [Sinha 96].

Ces expériences montrent bien l'apport des éléments externes du visage dans le processus de reconnaissance faciale humaine. Cependant, de nombreux traitements de reconnaissance faciale ne se basent que sur les éléments internes du visage. Ainsi l'apport de la chevelure pourrait amener des paramètres supplémentaires pour améliorer ces processus.

1.2.1 Identification & authentification de personne

Nous avons vu dans le paragraphe précédent le rôle de la chevelure dans le processus humain de reconnaissance faciale. La question qui se pose à nous maintenant est comment intégrer l'information portée par les cheveux dans les techniques d'identification de personne. Un processus basé uniquement sur l'analyse des paramètres capillaires est a priori inefficace au vu des résultats mis en évidence par Jarudi *et al* [Jarudi 07]. La principale différence entre l'identification et l'authentification d'une personne reste la connaissance ou non de la photo de référence pour la comparaison. Dans un cas on cherche à évaluer la similitude entre 2 photos, dans l'autre on cherche à maximiser la vraisemblance entre 2 photos. La chevelure est très facilement transformable, même si la majorité des personnes gardent une coiffure identique. La chevelure reste le paramètre facial le plus simple, rapide et peu coûteux à utiliser pour changer d'apparence.

Dans [Yacob 06], Yacob et Davis utilisent une caractérisation du masque capillaire pour améliorer une application de reconnaissance faciale par Eigenface. Suite à une segmentation du masque capillaire qui sera détaillée dans le chapitre suivant, 8 paramètres sont extraits de la chevelure : la couleur, la longueur, la texture, le volume, la surface, la symétrie, la position de la raie et la "hairline" ou ligne du crâne. Dans un premier temps, chaque paramètre a été testé individuellement pour un processus d'identification basé sur une analyse par composante principale. Les résultats obtenus montrent un très faible taux de reconnaissance avec moins de 5% pour les paramètres de la longueur, de la surface, de la symétrie, du volume et de la position de la raie. Un maximum de bonne reconnaissance est acquis pour la texture avec presque 20%. Ces

résultats confirment les hypothèses citées auparavant sur la faible performance des paramètres uniques pour la reconnaissance de personne.

Néanmoins, l'ajout des paramètres de la chevelure au système de reconnaissance par Eigenfaces permet une amélioration des résultats d'identification. Ces résultats sont d'autant plus intéressants car ils permettent d'améliorer la reconnaissance de visages pour des images dont les conditions d'acquisition ne sont pas contrôlées : changement d'illumination et visages occultés par des lunettes ou des écharpes/cols.

Ces expériences mettent bien en évidence le rôle de la chevelure pour l'amélioration des algorithmes de reconnaissance faciale. La chevelure n'est peut être pas un paramètre aussi utile dans un processus d'authentification que d'identification mais pourrait permettre la création d'image de référence pour améliorer la comparaison. Ainsi l'apport de la chevelure permettrait la création d'avatar, à partir d'une personne à identifier, sur lequel on pourrait faire varier les paramètres de la coiffure. Une collection de portrait robot pourrait être constituée afin d'estimer la nouvelle apparence d'une personne. Une autre approche pourrait être aussi utilisée dans les algorithmes de vieillissement de personnes, en effet la caractérisation de la chevelure pourrait permettre une meilleure estimation de l'évolution possible d'une chevelure au cours du temps.

Cette piste de développement doit s'appuyer sur la modélisation la plus réaliste possible d'une coiffure afin de correspondre le plus possible au visage étudié. Ainsi le développement de la segmentation capillaire pourrait aussi apporter des solutions pour améliorer ou pour optimiser des applications de synthèse d'images de chevelure.

1.3 La coiffure, reflet du réalisme

La chevelure a toujours fasciné de nombreux artistes et de nombreuses oeuvres lui sont dédiées. Les artistes de tout temps ont toujours voulu dépeindre cette matière qui propose tant de teintes et de reflets, tant de formes changeantes pouvant bouger au grès du vent. Au fil des siècles les représentations picturales des cheveux ont toujours été de plus en plus réalistes en partant du simple dessin, en passant par la peinture et pour finir aujourd'hui par des images de synthèse avec des rendus de qualité photographique (1.11)



FIGURE 1.11 – **Evolution de la représentation picturale de la chevelure** : (a) Détail de la Naissance de Venus par Botticelli (1485), (b) Dessin au fusain, (c) Dessin Manga, (d) Image de synthèse tirée du film Final Fantasy VII.

1.3.1 La Synthèse de visage

A l'instar des autres technologies, la synthèse d'images et l'animation d'objets virtuels ont connu un fort développement ces dernières années. L'augmentation de la puissance de calcul des ordinateurs a permis le développement des techniques de synthèse d'images que ce soit pour la création d'image, de photo ou de film. Si la synthèse et le rendu de scènes naturelles sont aujourd'hui des plus réalistes, l'animation d'avatars humains demeure une difficulté. En effet si les visages sont visuellement réalistes, la modélisation d'expressions, d'émotions ou l'animation de la parole restent très difficiles. L'animation de ces avatars reste trop figée ce qui perturbe la perception des visages ainsi modélisés.

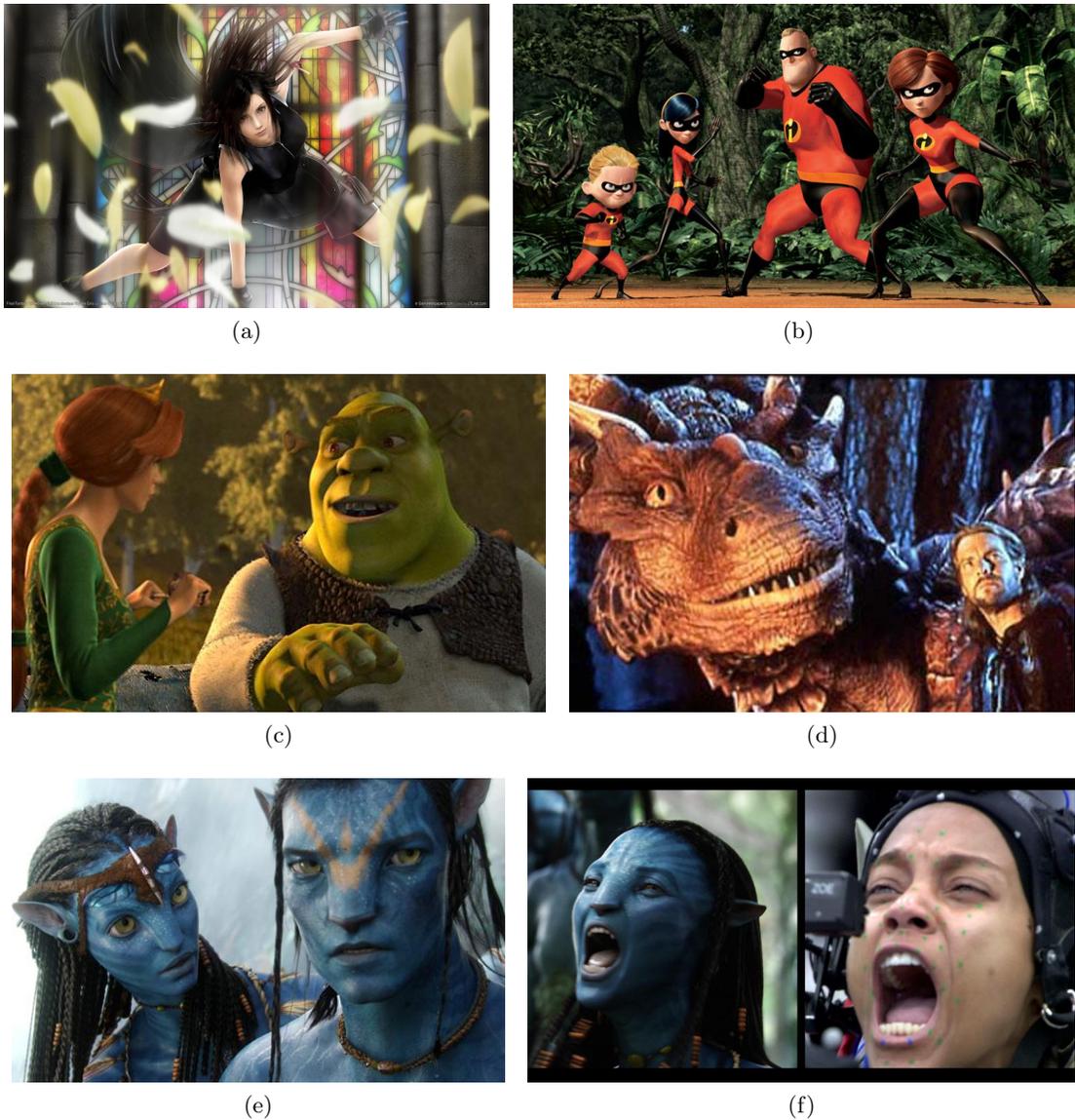


FIGURE 1.12 – **Film en images de synthèse** : (a) Final Fantasy : Advent Children , (b) Les Indestructibles, (c) Shrek, (d) Coeur de Dragon, (e) Avatar, (f) Avatar Processus de biométrie pour afficher les animations faciales de l'acteur.

Ces dernières années, la synthèse d'image a surtout été développée dans l'industrie cinématographique où des films entièrement en images de synthèse ont vu le jour : Final Fantasy (figure 1.12a), Les Indestructibles (figure 1.12b), Shrek (figure 1.12c), Avatar (figure 1.12e)...

De nombreux traitements ont été développés pour que l'animation d'avatars humains ou humanoïdes colle toujours au plus près à la réalité. Par exemple dans le film *Coeur de Dragon* (figure 1.12d), un morphing entre le visage du Dragon et le visage de Sean Connery qui lui prête sa voix, a été effectué afin que le dragon ait des expressions humaines. Dernièrement on ne peut que citer le film *Avatar* de James Cameron qui a repoussé encore plus loin les limites des films en images de synthèse. Outre le fait que le film soit diffusé en 3 dimensions, un système biométrique très évolué a été mis au point afin de rendre l'animation des avatars encore plus réaliste. En effet une caméra filmait continuellement le visage des acteurs auxquels on avait ajouté des points de références pour lier l'animation des avatars au jeu des acteurs (figure 1.12f).

L'utilisation d'avatars n'est pas réservée au seul monde du cinéma et de nombreuses applications d'interaction homme machine se munissent d'avatars 3D. On peut notamment citer le projet TELMA (TELEphonie à l'usage des MAalentendants), projet visant à donner l'accès à la téléphonie à des personnes malentendantes. Le but étant de modéliser le LPC (Langage Parlé Complété) qui associe des postures de la main et des doigts à la parole et la forme de la bouche. Le projet comporte les deux aspects de la communication (figure 1.13) : si la personne malentendante est le locuteur, une analyse de l'image devra déterminer les sons définis par la forme de la bouche et la position des mains afin de déterminer le message à transmettre pour une synthèse vocale. Inversement si la personne est malentendante, le message sonore reçu devra être analysé afin d'extraire les différents sons qui devront être traduit en LPC et reproduit par un avatar 3D animé.

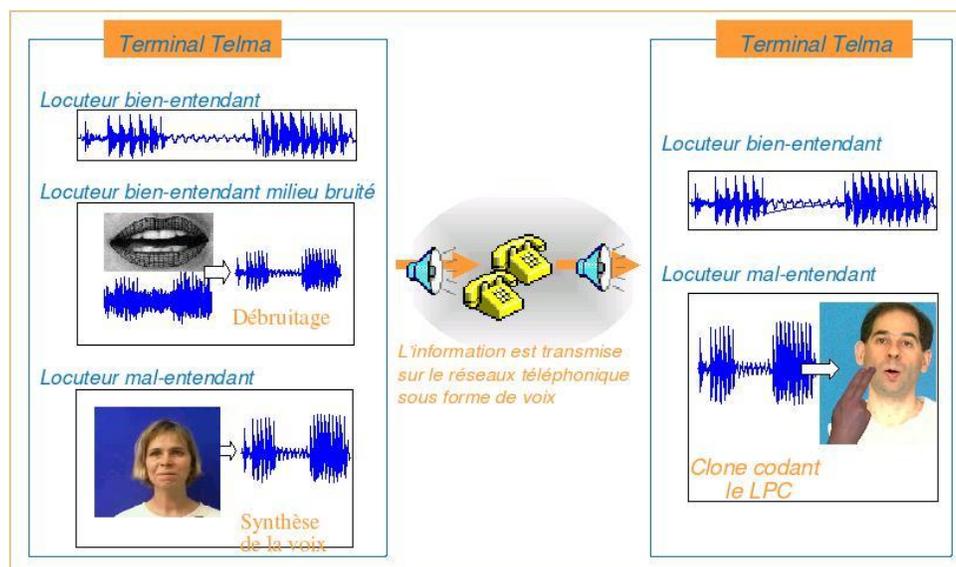


FIGURE 1.13 – **Projet TELMA** :Processus de communication du Projet TELMA

1.3.2 La Synthèse Capillaire

Étant donnée la multiplication des personnages virtuels notamment dans les jeux vidéo et les films d'animation, la simulation de cheveux fait l'objet d'un intérêt croissant de la part des spécialistes de la synthèse d'images. En effet le rendu de la synthèse de la chevelure joue un rôle important dans le rendu réaliste d'un visage tant dans la représentation que dans l'animation. Si le rendu de cheveux a aujourd'hui atteint un réalisme pratiquement parfait, il n'en est cependant pas de même pour l'animation car c'est l'une des tâches les plus difficiles et les plus longues à réaliser. Dans *Final Fantasy* (figure 1.12a), la simulation des 60 000 cheveux individuels du personnage principal avait mobilisé 20% du temps de calcul total de la production. Le lecteur intéressé trouvera plus d'information dans les travaux de thèse "Simulation de chevelures virtuelles" de Florence Bertails [Bertails 06].

La majorité des travaux de synthèse de la chevelure se base sur une modélisation à partir de l'étude physique d'un cheveu (*annexe A*) et non sur une analyse des caractéristiques d'un cheveu dans une image.

On peut néanmoins citer les travaux de Chen *et al* pour une application avec des images 2D [Chen 06]. Ils présentent une technique de synthèse capillaire à partir d'une coiffure réelle et d'un sketch³. Une chevelure est segmentée manuellement et sert de modèle à leur masque synthétique qui devra être une copie la plus fidèle possible de la chevelure initiale. Leurs travaux sont d'autant plus intéressants que leur synthèse se base sur l'analyse d'une image à partir d'une chevelure réelle et donc se base sur la caractérisation de la chevelure qui pourrait être utile pour l'analyse et la segmentation capillaire.

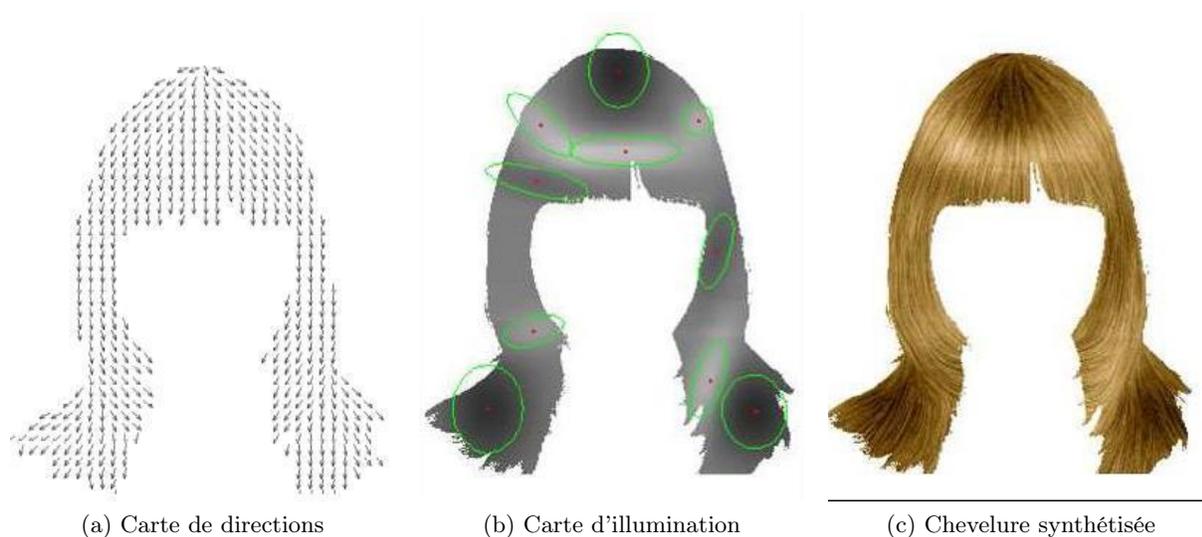


FIGURE 1.14 – **A Generative Sketch Model for Human Hair Analysis and Synthesis** : Exemple de coiffure synthétisée (c) à partir de la décomposition de 2 différentes cartes de caractérisation d'une chevelure : (a) la carte de direction, (b) la carte d'illumination ainsi que d'un modèle couleur (blond dans l'exemple). [Chen 06]

3. Un sketch est un dessin rapide fait à main levée. Il sert généralement de croquis rapide pour une réutilisation future

1.4 Conclusion

Dans ce chapitre, nous avons présenté l'intérêt et le rôle que peut apporter l'analyse et la segmentation de la chevelure pour divers types d'applications :

- **Maquillage virtuel** : L'analyse et la segmentation de la chevelure pourrait permettre l'amélioration de ces applications en proposant une possibilité supplémentaire de maquillage, en adaptant au visage les chevelures synthétiques et en améliorant le rendu visuel des applications de colorations capillaires.
- **Application de reconnaissance faciale** : Élément pour la reconnaissance de personnes, l'analyse capillaire pourrait apporter des pistes d'améliorations. Dans les applications biométriques de reconnaissances faciales, elle apporterait plusieurs paramètres caractéristiques.
- **Synthèse d'images** : Avec l'explosion de l'utilisation des images de synthèses, la représentation de la chevelure est un élément primordiale du rendu réaliste d'un visage virtuel. L'analyse capillaire pourrait apporter de l'information pour améliorer ou guider la synthèse de chevelures.

La chevelure est donc bien un facteur d'information qui peut apporter de réelles solutions pour améliorer ou diversifier ces applications. Elle est pourtant peu utilisée et son analyse est encore très peu traitée dans le domaine du traitement d'images.

2 La Segmentation Capillaire



Sommaire

2.1	État de l'Art sur l'Analyse Capillaire	21
2.1.1	La Chevelure : un élément à part entière du visage	21
2.1.2	Travaux fondateurs de la segmentation capillaire	23
2.1.3	Approche combinant une segmentation couleur de peau et des cheveux	25
2.1.4	Segmentation par analyse couleur et estimation de la localisation	27
2.1.5	Segmentation de la chevelure par approche bayésienne et Meanshift	29
2.1.6	Caractérisation de la chevelure pour sa synthèse	31
2.1.7	Discussion Générale sur la segmentation capillaire	32
2.2	Techniques de Segmentation	34
2.2.1	L'approche "Contours"	34
2.2.1.1	Les opérateurs dérivatifs	34
2.2.1.2	Les contours actifs	35
2.2.1.3	Analyse des approches "contours"	35
2.2.2	Approche Régions	37
2.2.2.1	Les méthodes par agrégation	37
2.2.2.2	Les méthodes de division-fusion	38
2.2.2.3	Les méthodes de modélisation statistique	38
2.2.2.4	Analyse des approches "régions"	38
2.3	Conclusion	39

Dans ce chapitre, nous présenterons un état de l'art des travaux d'analyse et de segmentation capillaire (*section 2.1*), puis une revue non exhaustive des techniques de segmentation qui pourraient être utilisées pour la segmentation capillaire. Dans un premier temps, nous étudierons les techniques dites par approches "*contours*" (*section 2.2.1*) et dans un second temps, les techniques dites par approches "*régions*" (*section 2.2.2*).

2.1 État de l'Art sur l'Analyse Capillaire

La chevelure a longtemps été considérée comme un paramètre instable à cause de sa diversité de représentation de sa facilité et de sa rapidité à pouvoir changer son apparence. De ce fait, elle est peu traitée dans le domaine du traitement d'images que ce soit pour des applications de segmentations, de biométries ou de synthèse d'images. C'est récemment que l'intérêt pour la chevelure est apparu avec des premières publications en synthèse d'images à partir de 1989 et seulement à partir de 1998 pour les publications en traitement d'images [Kampmann 98]. À partir du début des années 2000, la chevelure est particulièrement étudiée en synthèse d'images alors qu'il faut attendre le milieu des années 2000 pour commencer à trouver des publications en traitement d'images avec un article fondamental en 2006 [Yacoob 06] qui pose les fondements de la segmentation capillaire. Article de référence pour les différents travaux actuels [Rousset 08] [Zhang 08] [Lee 08] [Wang 09].

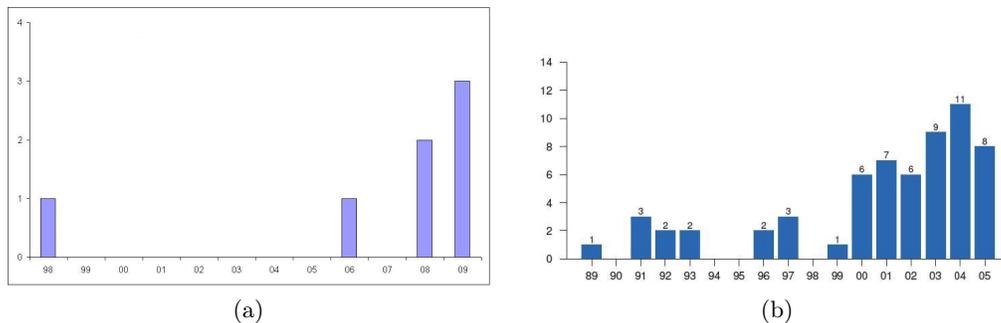


FIGURE 2.1 – **Évolution du nombre de publication sur la chevelure** : Nombre de publications sur la segmentation capillaire, par an, (b) Nombre de publications sur la synthèse de cheveux, par an [Bertails 06].

Nous allons présenter dans la suite de ce chapitre, l'ensemble des publications traitant de l'analyse et/ou de la segmentation capillaire que nous ayons pu trouver dans la littérature.

2.1.1 La Chevelure : un élément à part entière du visage

Les premiers travaux que nous ayons trouvé traitant de la chevelure datent de 1998 [Kampmann 98]. Dans cette publication, Kampmann propose une approche pour optimiser la compression de vidéo contenant des visages. Il décrit ainsi un algorithme de segmentation automatique du visage en différentes parties pour compresser chaque partie en fonction de l'information qu'elle contient. Kampmann estime que l'information portée par la face est plus importante que celle portée par les autres parties du visage telles que les oreilles, le cou et la chevelure.

La première étape de l'algorithme consiste à effectuer une segmentation couleur de la peau. Un modèle couleur de peau est calculé à partir de fenêtres échantillons situées sur les joues. Elles

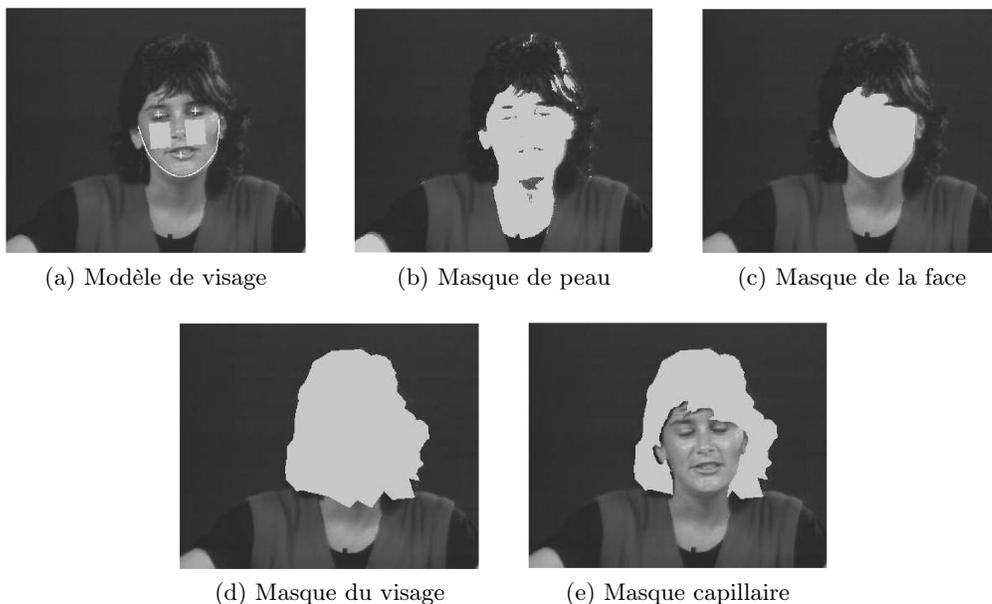


FIGURE 2.2 – **Segmentation of a Head into Face, Ears, Neck and Hair for Knowledge Based Analysis Synthesis Coding of Videophone Sequences** : (a) Représentation du modèle de localisation du visage (yeux, bouche, arc mandibulaire et fenêtres échantillons de peau), (b) Image de la segmentation de peau, (c) Image résultat de la définition de la zone face, (d) Image résultat de la segmentation du visage, (e) Image résultat de la définition de la zone capillaire. [Kampmann 98]

sont localisées grâce à une détection de la position des yeux, de la bouche et de l’arc mandibulaire (*figure 2.2a*) [Kampmann 97a]. Le masque de peau ainsi obtenu (*figure 2.2b*) est divisé en trois parties : la face, les oreilles et le cou. La face étant définie comme la zone appartenant au masque de peau comprise dans la zone à l’intérieur de l’arc mandibulaire. Tous les artefacts pouvant exister dans le modèle couleur tels que les yeux ou les sourcils sont comblés par un traitement classique de fermeture de l’image (*figure 2.2c*). Les oreilles sont définies, quant à elles, à partir du masque de peau et de visage, ce sont les zones latérales au masque de visage appartenant au masque de peau. Enfin le cou est défini comme la partie inférieure du masque de peau se trouvant en dessous de l’arc mandibulaire. Afin de limiter sa taille, un modèle de longueur moyenne de cou est estimé à partir des informations de distance entre les yeux et la bouche.

Finalement, suite à une détection de la zone visage grâce à un algorithme de suivi de visage dans une séquence vidéo (*figure 2.2d*) [Kampmann 97b], la zone du masque capillaire (*figure 2.2e*) est estimée en soustrayant du masque visages les zones déjà segmentées de la face, des oreilles et du cou.

Ces travaux constituent donc les premiers sur la segmentation capillaire. Cette étude n’est pas directement liée à une analyse et une caractérisation des paramètres capillaires dans une image. Néanmoins, il obtient des résultats intéressants mais dépendants beaucoup de son algorithme de détection de l’ensemble du visage, en effet la chevelure dépend directement de la zone segmentée. Cet article est surtout intéressant car il montre que la segmentation capillaire est étroitement liée à une analyse et une segmentation du visage. Ainsi la segmentation du masque de peau est

très utile car elle permet de définir une partie de la frontière naturelle de la chevelure. Même si la localisation précise de la chevelure n'est pas simple à définir, nous pouvons facilement admettre qu'elle se trouve autour du visage. De plus bien que cette approche soit assez simple, avec des conditions de segmentation contrôlées, Kampmann a déjà été confronté à des problèmes de segmentation de la chevelure au niveau de la frontière entre la zone de visage détectée et le reste du corps en particulier pour les chevelures longues.

2.1.2 Travaux fondateurs de la segmentation capillaire

L'article que l'on peut citer comme référence dans le domaine de la segmentation capillaire est très récent puisqu'il ne date que de 2006, mais il définit un cadre pour les travaux actuels traitant du sujet [Yacoob 06]. C'est d'ailleurs à partir de ces travaux que nous avons commencé notre étude. Dans cette publication, Yacoob *et al* présentent la première approche de segmentation capillaire utilisant des paramètres caractéristiques de la chevelure. Pour leur étude, ils limitent leur base d'images de traitement pour des modèles vus de face.

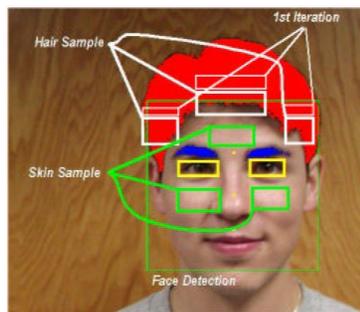


FIGURE 2.3 – "**Detection and Analysis of Hair**" - **Modèle de la segmentation couleur** : Localisation des fenêtres échantillons pour la définition des modèles couleurs. En vert les fenêtres échantillons pour la peau et en blanc celles pour les cheveux. [Yacoob 06]

Leur algorithme de segmentation se divise en 4 étapes :

- **La détection du visage** : Comme pour Kampmann [Kampmann 98], le visage définit le point d'initialisation de leur algorithme. Ils emploient l'algorithme de détection de visage en cascade de la librairie OpenCV. Basé sur la méthode Adaboost, ce détecteur est développé sur les travaux de Viola *et al* [Viola 01][Viola 04] ainsi que son amélioration proposée par Lienart *et al* [Lienhart 03].
- **La détection des yeux** : une fois le visage détecté, ils localisent les yeux avec un détecteur basé lui aussi sur l'algorithme en cascade de la librairie OpenCV. La détection couplée du visage et des yeux leur permet de normaliser la taille du visage dans l'image afin de pouvoir comparer les différents masques capillaires obtenus.
- **La segmentation du masque de peau** : le masque de peau est obtenu par une segmentation couleur. Un modèle colorimétrique de peau est calculé à partir de l'étude de 3 fenêtres échantillons situées sur des zones homogènes du visage, une sur le front et les 2 autres sur les joues (*figure 2.3*). La segmentation couleur est effectuée dans l'espace RGB par un algorithme de minimisation de la distorsion couleur. Cet algorithme est détaillé plus en détails dans la section 5.3.

- **La segmentation du masque capillaire** : Le masque de la chevelure est aussi obtenu par une segmentation couleur basée aussi sur l'algorithme de minimisation de la distorsion couleur. Contrairement à la segmentation de la peau, plusieurs modèles colorimétriques sont calculés à partir de 3 fenêtres échantillons : une au sommet du crâne et une de chaque côté de la partie supérieure du visage (*figure 2.3*). En effet la présence de pixels cheveux dans ces zones n'est pas obligatoire. Dans un premier temps, pour chaque modèle couleur, les pixels proches du modèle couleur de peau sont retirés. Puis dans un second temps, les 3 modèles couleurs sont comparés : si la distance couleur (ie la distance entre la couleur moyenne de chaque modèle dans l'espace RGB) est faible, alors un modèle couleur globale est recalculé sinon le modèle couleur de la fenêtre échantillon du sommet de crâne est pris comme référence. La segmentation du masque capillaire se fait itérativement à partir des fenêtres échantillons. À chaque itération, les fenêtres sont agrandies et un nouveau modèle couleur est calculé. Si le modèle courant est proche du modèle couleur de référence on passe à l'itération suivante en considérant le modèle courant comme référence sinon on stoppe la segmentation. Un fois le masque capillaire obtenu, il est nettoyé par des algorithmes classiques de traitement d'images afin d'obtenir un masque d'une seule composante connexe.

Ces travaux présentent réellement la première approche de segmentation capillaire à partir de l'analyse d'un paramètre caractéristique de la chevelure dans une image. Bien que l'évaluation des résultats soit validée subjectivement par validation visuelle d'un opérateur, l'approche mise en place donne de bons résultats avec 71% de segmentation estimées réussies sur une base d'images où les conditions d'acquisitions étaient peu contrôlées et jusqu'à 93% de bonnes segmentations estimées sur des images de la base AR [ARb]. Ces résultats montrent que la couleur de la chevelure est un paramètre assez discriminant pour la segmentation capillaire.

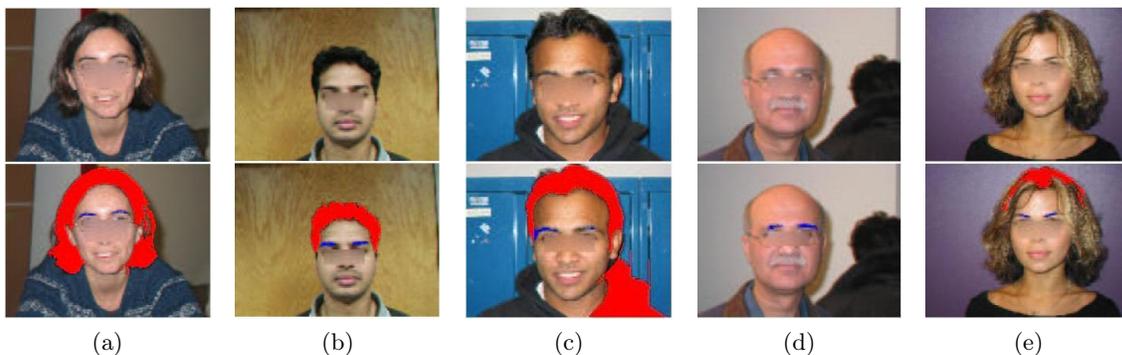


FIGURE 2.4 – "**Detection and Analysis of Hair**" - Exemple de Segmentation : (a) Bonne segmentation, (b) Bonne segmentation, (c) Mauvaise segmentation : une partie du pull a été sur segmenté, (d) Mauvaise segmentation : une personne chauve, (e) Mauvaise segmentation : différence de couleur entre les racines et le reste de la chevelure. [Yacoob 06]

Cette première approche met aussi en évidence certaines limites propres à la segmentation capillaire :

- La sur segmentation de pixels de vêtements (*figure 2.4c*). La définition d'un seul et unique descripteur de la chevelure est insuffisant pour des éléments de couleur similaire à la chevelure. Une piste d'amélioration serait de définir au moins un second paramètre pour

améliorer la segmentation.

- L'initialisation du modèle couleur qui a échoué sur une personne chauve (*figure 2.4d*). La définition des zones de référence de chevelure n'est pas simple car la répartition d'une chevelure d'une personne à une autre n'est pas stable. Une piste d'amélioration serait de définir un indice de fiabilité des pixels présents dans les fenêtres échantillons, indice lié à une autre caractéristique de la chevelure par exemple.
- La sous-segmentation de la chevelure à cause d'une différence de couleur entre la couleur des racines et celles des pointes d'une même chevelure (*figure 2.4e*). La définition d'un modèle de couleur unique pour une chevelure n'est pas simple et la définition d'une seule zone d'étude n'est pas forcément la plus représentative de l'ensemble de la chevelure.

Dans la seconde partie de leur publication, Yacoob *et al* ont défini 8 paramètres pour caractériser le masque capillaire segmenté : la couleur, la longueur, la texture, le volume, la surface, la symétrie, la position de la raie et la "hairline" ou ligne du crâne. Ils ont testé individuellement chaque paramètre dans un processus de reconnaissance de personne. La texture de la chevelure permet à elle seule presque 20% de bonnes reconnaissances alors que les autres paramètres ne permettent environ que 5% de bonnes reconnaissances. Ces résultats montrent que la texture pourrait contenir assez d'information utile pour aider à la segmentation.

2.1.3 Approche combinant une segmentation couleur de peau et des cheveux

Dans cette publication Zhang *et al* proposent une nouvelle approche pour la détection de visage en combinant une segmentation couleur de la chevelure et de la peau [Zhang 08] [Zhang 09]. Cette prise en compte de la chevelure vise à améliorer la détection d'un visage dans une image quelle que soit son orientation (de face, de côté ou de dos ainsi que pour des prises de vue de dessus ou de dessous).

Leur approche se base sur l'étude conjointe des espaces couleurs XYZ et HSV par la modélisation de 2 classifieurs couleurs, un pour la peau et un pour la chevelure. Ils ont choisi un classifieur couleur représentant la densité de probabilité d'appartenir au modèle couleur par un mélange de gaussienne estimé par l'algorithme d'*expectation-maximization* (EM). Les 2 modèles couleurs sont calculés à partir de bases d'images d'échantillons segmentés manuellement (*figure 2.5*).



FIGURE 2.5 – An Accurate Algorithm for Head Detection - Échantillons couleurs : 1^{ère} ligne échantillons couleurs de peau, de la 2^{ème} à la 5^{ème} lignes échantillons couleurs de chevelure : brun, roux, blond et noir. [Zhang 08]

- **Modèle de couleur de peau** : Une fois le modèle calculé, la probabilité pour un pixel x d'être un pixel de peau S peut s'exprimer pour chaque espace couleur respectivement :

$$\begin{aligned} p^{XYZ}(x|S) &= p(x_X|S) \cdot p(x_Y|S) \cdot p(x_Z|S) \\ p^{HSV}(x|S) &= p(x_H|S) \cdot p(x_S|S) \cdot p(x_V|S) \end{aligned} \quad (2.1)$$

et la probabilité combinée est estimée selon l'équation :

$$P_{peau}(x) = \delta_1 p^{XYZ}(x|S) + \delta_2 p^{HSV}(x|S) \quad (2.2)$$

avec les pondérations δ_1 et δ_2 déterminées expérimentalement à partir de la base d'entraînement.

Finalement le masque de localisation des pixels de peau est obtenu en seuillant la fonction de densité de probabilité P_{peau} selon un seuil th_{peau} déterminé expérimentalement. Pour optimiser leurs résultats, ils ont combiné leur approche avec un autre algorithme de détection de pixel de peau [Hsu 02]. Au final leur masque de localisation de pixels de peau est la combinaison logique de ces 2 masques.

- **Modèle de couleur de la chevelure** : Pour prendre en compte la diversité de coloration d'une chevelure, 4 catégories de couleurs représentatives des teintes de cheveux ont été définies : noir, blond, brun et roux. Pour chaque catégorie, une base d'entraînement est utilisée pour estimer les fonctions de densité de probabilité $p^{XYZ}(x|G)$ et $p^{HSV}(x|G)$, avec $G \in \{noir, blond, brun, roux\}$ représentant les différents états possibles du pixels.

La probabilité combinée est estimée selon l'équation

$$P_{cheveu}(x) = \omega_1 p^{XYZ}(x|G) + \omega_2 p^{HSV}(x|G) \quad (2.3)$$

avec les pondérations ω_1 et ω_2 déterminés expérimentalement à partir des bases d'entraînement. Au final la probabilité pour un pixel x d'être un pixel cheveux est estimée selon l'équation :

$$P_{cheveu}(x) = \arg \max_G (P_{cheveu}(x|G)). \quad (2.4)$$

Finalement le masque de localisation des pixels de cheveu est obtenu en seuillant la nouvelle fonction de densité de probabilité P_{cheveu} selon un seuil th_{cheveu} déterminé expérimentalement.

Une fois les 2 masques obtenus, la carte de localisation du visage est créée par l'union des masques de peau et de chevelure. Après avoir procédé à une fermeture, un étiquetage est effectué afin d'identifier les différentes composantes connexes détectées. Pour chaque composante, une ellipse englobante est calculée ainsi que plusieurs paramètres tels que son excentricité ou sa surface (*figure 2.6a*). Ces paramètres sont comparés à des conditions de taille et d'orientation pour valider le visage trouvé.

Finalement, ils évaluent leur méthode sur 4 différentes bases d'images, 3 représentants a priori un unique modèle dans une image et une comportant plusieurs visages. Ils obtiennent le résultat parfait de 100% de bonnes détections de visages et aucune fausse alarme. Cependant, aux vues des images résultats publiées, on peut émettre un léger *bémol* à ces résultats, la plupart des images représentant une unique personne devant un fond uniforme (*figure 2.6b*).

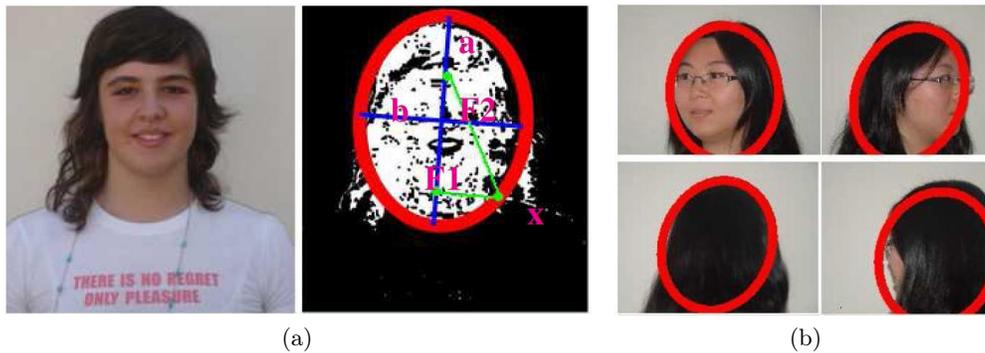


FIGURE 2.6 – **An Accurate Algorithm for Head Detection - Résultats** : (a) Exemple de carte de localisation du visage avec la représentation de l'ellipse englobante correspondant au visage détecté, (b) Exemple de bonnes détections. [Zhang 08]

Néanmoins, cette nouvelle approche de détection du visage propose une alternative pour la segmentation couleur de la chevelure en se basant sur des modèles couleurs définissant des couleurs types. Mais la segmentation précise du masque capillaire n'étant pas la finalité des travaux, l'évaluation de leur résultat ne reste que partielle. D'autant plus que la définition de l'ellipse englobante permet de s'affranchir des éventuelles erreurs de segmentation.

2.1.4 Segmentation par analyse couleur et estimation de la localisation

Dans ce papier, Lee *et al* proposent une nouvelle approche pour la segmentation du visage et de la chevelure basée sur une étude de la couleur de peau et de la chevelure ainsi que sur une estimation de la localisation de la chevelure par rapport au visage [Lee 08].

Leur algorithme de segmentation commence par la résolution de l'équation $S(x) = f_x$ qui à tout pixel x associe une étiquette $f_x \in 0, 1, 2$ avec 0 pour un pixel du fond, 1 pour un pixel de cheveu et 2 pour un pixel du visage. L'étiquetage de $S(x)$ sera guidé par les distributions de probabilité $P(x|f)$ pour le modèle de localisation et $P(z|f)$ pour le modèle couleur où z représente la valeur du pixel dans l'espace RGB ou La^*b^* .

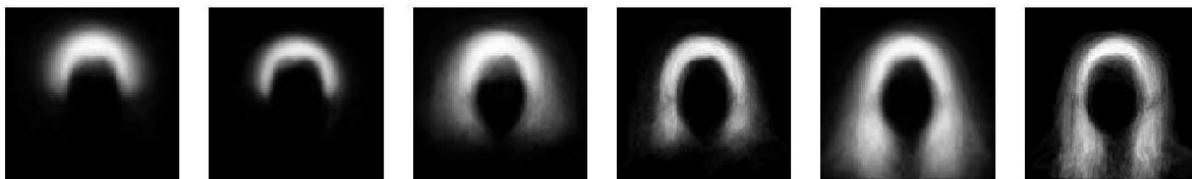


FIGURE 2.7 – **Distribution de densité de probabilité de la chevelure** : Représentation des 6 types de coiffures pour des cheveux de plus en plus longs (gauche à droite) [Lee 08]

- **La densité de probabilité de localisation $P(x|f)$** : les distributions de probabilité de la localisation des pixels de visage et de la chevelure sont modélisées à partir de vérités terrains obtenues pour des bases d'apprentissage segmentées manuellement. Pour prendre en compte la variabilité de la chevelure, les modèles de chevelure ont été séparés en 6 classes (*figure 2.7*). Pour chaque classe (1 de visage et 6 de chevelure), la distribution de

probabilité $P(x|f)$ est calculée simplement en comptant pour chaque position x le nombre de pixels appartenant à chaque classe f_x .

- **La densité de probabilité de distribution couleur** $P(z|f)$: la modélisation de la distribution de probabilité $P(z|f)$ des modèles couleurs s'effectue en 2 étapes : la première 'offline' à partir de bases d'apprentissage et la seconde 'online' à partir de l'étude de l'image à segmenter.
- Etape 'offline' : les modèles couleurs 'offline' sont définis comme pour [Zhang 08] par une fonction de densité de probabilité représentée par un mélange de gaussiennes dont les paramètres sont estimés grâce l'algorithme d'*expectation-maximization*(EM). Ainsi les distributions $P(z|f = 1)$ et $P(z|f = 2)$ sont calculées sur les mêmes bases de référence que pour la localisation. Les pixels cheveux sont cette fois ci séparés en 5 catégories de couleur : noir, brun, blond, roux et gris. Finalement la distribution de probabilité pour les pixels du fond $P(z|f = 0)$ est initialisée comme une distribution uniforme.
- Etape 'online' : Les modèles couleurs 'online' sont mis à jour à partir de l'étude de l'image à traiter, les pixels sont prélevés dans les zones de maximum de probabilité obtenues à partir de la distribution de localisation $P(x|f)$. La distribution $P(z|f = 0)$ est mis à jour par le calcul du mélange de gaussiennes des pixels du fond. Les pixels de visage ne correspondant pas au modèle couleur 'offline' ne sont pas pris en compte et la distribution $P(z|f = 2)$ est recalculée à partir des pixels restants. La mise à jour de la distribution des pixels cheveu $P(z|f = 1)$ est similaire, après avoir effectué le choix du modèle couleur à partir du maximum de probabilité correspondant pour chaque classe.

Une fois les modèles de distribution de probabilité calculés, la segmentation est soit opérée par une approche par Loopy Belief Propagation [Yedidi 03] liée à l'utilisation de Champs de Markov, soit par une approche Graph Cut [Boykov 01].

Lee *et al* utilise l'approche Graph Cut basée sur les travaux de Boykov *et al* [Boykov 01]. Le processus de segmentation revient alors à un problème de minimisation d'une fonction d'énergie $E(S)$ qui mesure la qualité de la segmentation S . L'équation de E s'écrit sous la forme suivante :

$$E(S) = R(S) + \alpha B(S). \quad (2.5)$$

avec α qui définit la pondération entre les 2 fonctions d'énergie.

La fonction R mesure la qualité de segmentation des pixels correspondant aux paramètres capillaires de la distribution $P(z|f)$ de couleur et $P(x|f)$ de localisation défini précédemment :

$$R(S) = \sum -\log P(z|f) - \beta \log P(x|f) \quad (2.6)$$

avec β le coefficient de pondération relatif entre le terme de probabilité couleur et de probabilité de localisation.

La fonction B prend en compte l'homogénéité du voisinage local et s'exprime selon l'équation :

$$B(S) = \sum_{x_i, x_j} \delta(S(x_i) \neq S(x_j)) \exp(-\gamma \|z_i - z_j\|^2) \quad (2.7)$$

avec $\delta(S(x_i) \neq S(x_j)) = 1$ si les pixels i et j ont la même étiquette f et 0 sinon. γ est une constante normalisée qui peut être calculée comme l'inverse du contraste moyen des pixels voisins.

Ainsi, étant donnée une image à traiter, le processus revient à trouver la segmentation S^* minimisant la fonction d'énergie E :

$$S^* = \arg \min_S E(S) \quad (2.8)$$

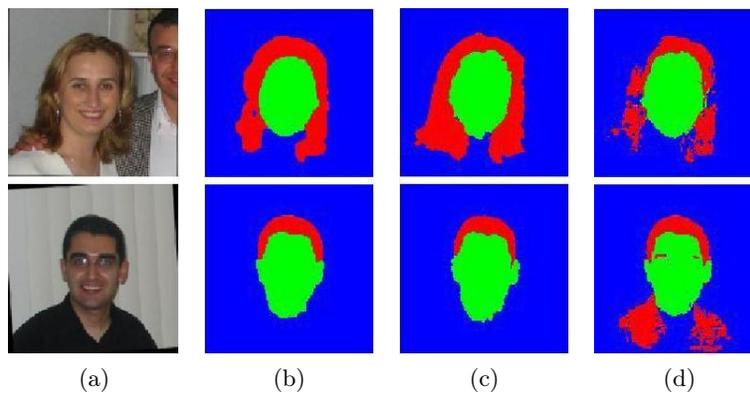


FIGURE 2.8 – **Markov random field models for hair and face segmentation** : Résultat obtenu pour les 2 approches (a) image originale, (b) Approche Graph Cut, (c) Approche Loopy Belief Propagation (d) comparaison avec la méthode couleur de Yacoob & Davis [Lee 08]

Au final, ils comparent leurs 2 méthodes avec une implémentation de l'approche couleur développée par Yacoob *et al* [Yacoob 06] adaptée à l'étiquetage des 3 classes (visage, chevelure et fond). Ils obtiennent un taux d'erreur de pixels mal étiquetés de 8.8% pour l'approche Graph Cut, de 9.4% par l'approche Loopy belief propagation et de 38.4% par l'approche couleur par minimisation de la distorsion couleur. Ces résultats peuvent s'expliquer en partie par le fait que la segmentation par minimisation de la distorsion couleur ne s'appuie que sur le paramètre couleur. Comme nous pouvons le voir sur l'exemple de la figure 2.8, dans le premier cas il y a sous segmentation de la chevelure blonde en partie à cause des différences de teinte entre les zones avec ou sans reflets. Dans le second cas, il y a sur-segmentation des pixels de vêtements car ils sont de couleurs proches de la chevelure. Ces exemples mettent en évidence certaines des limites expliquées à la section 2.1.2 et illustrent l'intérêt de ne pas se limiter qu'au seul paramètre couleur pour segmenter une chevelure. La localisation de la chevelure est donc un paramètre utile pour améliorer sa segmentation. Cependant l'approche par apprentissage développée dans l'article contraint les images traitées afin qu'elles correspondent aux modèles de localisation.

2.1.5 Segmentation de la chevelure par approche bayésienne et Meanshift

Dans cette publication Wang *et al* présentent une nouvelle approche de segmentation capillaire utilisant une segmentation par l'algorithme Mean Shift [Comaniciu 02] guidé par la probabilité d'occurrence ou localisation de la chevelure (HOPP) et un modèle couleur (GHCM) [Wang 09].

L'initialisation de leur algorithme est très similaire aux travaux de Lee *et al* [Lee 08] par l'utilisation des 2 mêmes paramètres pour décrire la position de la chevelure.

- **Le paramètre d'occurrence** (HOPP) : peut être comparé au paramètre de localisation [Lee 08] bien qu'une seule et unique classe ne soit définie. Il est calculé en comptant pour chaque position x le nombre de pixels appartenant au modèle de coiffure.
- **Le modèle couleur** (GHCM) : Leur modèle couleur est aussi modélisé grâce à un mélange de gaussiennes entraînée à partir d'une base de référence segmentée manuellement. Par contre un unique modèle couleur est défini contrairement à Zhang *et al* [Zhang 08] ou Lee *et al* [Lee 08] qui avaient défini des classes en fonction des teintes types d'une chevelure.

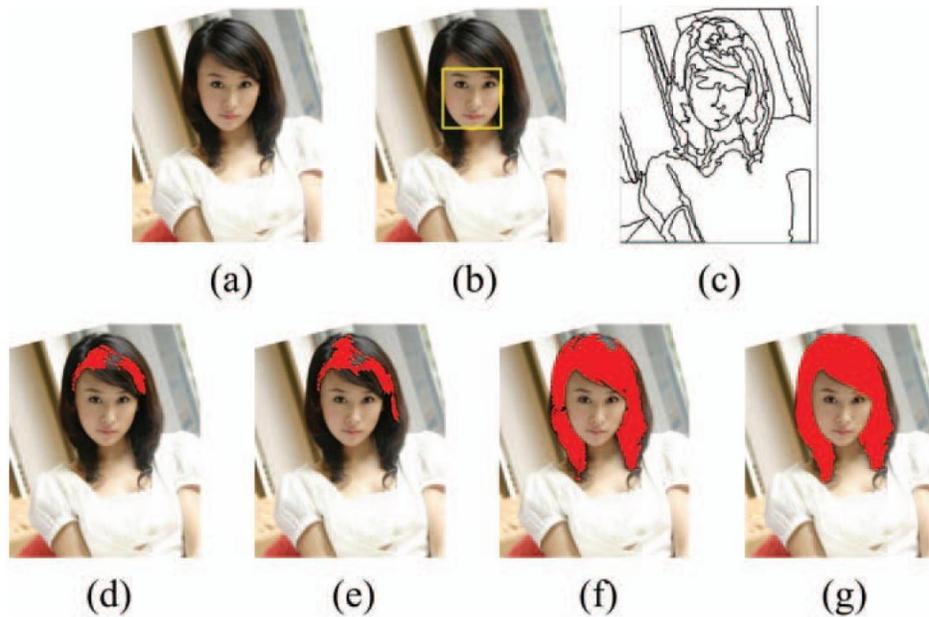


FIGURE 2.9 – **A Novel Two-tier Bayesian Based Method For Hair Segmentation** : (a) Image initiale, (b) Détection du visage, (c) Segmentation initiale par Mean Shift, (d) Marquage initial, (e) Propagation des germes à partir de la segmentation Mean Shift, (f) Classification des pixels cheveux, (g) Propagation de la classification à partir de la segmentation Mean Shift. [Wang 09]

Leur algorithme de segmentation capillaire se divise en 2 étapes. La première est initialisée par la définition de germes placés suivant une approche bayésienne combinant les 2 paramètres HOPP et GHCM. Pour chaque pixel (x, y) , on calcule la probabilité a posteriori d'être un pixel germe selon l'équation.

$$P(\text{Seed}|f(x, y)) = \frac{P(f(x, y)|\text{Seed})P(\text{Seed})}{P(f(x, y))} \quad (2.9)$$

avec $f(x, y)$ la valeur couleur du pixel, la probabilité conditionnelle $P(f(x, y)|\text{Seed})$ assimilée au modèle GHMC et la probabilité a priori $P(\text{Seed})$ représentée par le modèle HOPP.

Les germes sont définis suite à une pré segmentation de l'image par l'algorithme de Mean shift [Comaniciu 02] qui divise une image en différentes régions. Chaque région de la décomposition peut être considérée comme une région appartenant à la chevelure ou comme une région n'appartenant pas à la chevelure. Ainsi les germes retenus correspondent aux régions dont la proportion de pixels classifiés comme cheveux par le modèle bayésien est supérieure à un certain seuil.

Une fois les germes obtenus, un nouveau modèle couleur est calculé à partir de la distribution couleur des germes et un nouveau modèle bayésien est calculé

$$P(Hair|f(x, y)) = \frac{P(f(x, y)|Hair)P(Hair)}{P(f(x, y))} \quad (2.10)$$

avec $P(f(x, y)|Hair)$ le nouveau modèle gaussien de couleur.

Finalement la classification des pixels cheveux est obtenue selon l'équation :

$$C(x, y) = \begin{cases} 1 & \text{si } P(Hair|f(x, y)) > \delta_{Hair} \\ 0 & \text{sinon} \end{cases} \quad (2.11)$$

avec un pixel considéré comme un pixel cheveu si $C(x, y) = 1$ et δ_{Hair} un seuil défini empiriquement.

Au final, ils comparent leurs résultats avec nos résultats [Rousset 08] en comparant le taux de faux rejet et de fausse acceptation (cf section 3.5.2) sur la base d'images Caltech (cf section 3.4.3). Ils obtiennent des résultats similaires aux nôtres avec un taux de faux rejet de 3.7% et un taux de fausse acceptation de 36.6%. Ils ont aussi évalué leur algorithme sur une base plus importante de 1000 images et obtiennent des résultats du même ordre avec un Taux de faux rejet de 11.3% et un taux de fausse acceptation de 32.6%. On peut observer que leur approche est stable et donne des résultats intéressants avec un faible taux de faux rejet impliquant qu'en général l'ensemble de la chevelure est au moins segmentée. Comme pour les travaux de Lee [Lee 08], on peut conclure que l'ajout d'un second paramètre pour caractériser la chevelure, ici le paramètre d'occurrence (HOPP), a permis d'améliorer la segmentation capillaire. Mais pour les mêmes raisons, la définition par apprentissage de la localisation de la chevelure limitent le nombre d'images pouvant être traitées.

2.1.6 Caractérisation de la chevelure pour sa synthèse

Dans notre étude sur les travaux existants sur la segmentation capillaire, nous présenterons aussi les travaux de Chen *et al* [Chen 06]. Ils présentent une application de synthèse capillaire 2D à partir d'une chevelure pré segmentée. Bien que dans cet article, la chevelure soit segmentée manuellement, l'analyse capillaire utilisée pour caractériser leur modèle de synthèse peut aider à définir des paramètres qui pourraient être utiles à la localisation de la chevelure dans une image.

Ils étudient la chevelure dans l'espace couleur LUV et définissent ainsi 3 composantes principales : une carte de couleur, une carte de répartition de la lumière et une carte de texture.

- **La carte couleur** : elle est obtenue en mappant une couleur sur une carte d'intensité en niveaux de gris [0, 255]. La table des couleurs à mapper étant obtenue à partir de la discrétisation des composantes U et V.

- **La carte de répartition de la lumière** : elle est calculée à partir de la carte de luminance L . La carte de répartition de la lumière est définie après un filtrage gaussien passe bas pour obtenir la carte d'illumination et d'ombrage de la chevelure. Cette carte est représentée comme la superposition linéaire de gaussiennes. Pour chaque gaussienne, une ellipse englobante est calculée selon l'équation (figure 1.14) :

$$I_L(x, y) = \mu + \sum_i \alpha_i G(x - x_i, y - y_i, \theta_i, \sigma_{x_i}, \sigma_{y_i}) \quad (2.12)$$

avec μ la luminance moyenne, α un coefficient positif pour les illuminations et négatif pour les zones d'ombres et $x_i, y_i, \theta_i, \sigma_{x_i}, \sigma_{y_i}$ les paramètres de la $i^{\text{ème}}$ Gaussienne .

- **La carte de texture** : elle est calculée à partir de la carte de haute fréquence de la luminance. Elle est divisée en 3 niveaux : la texture apparente I_H , un champ de vecteur $V = (\theta, \phi, d_V)$ représentant respectivement l'orientation θ , l'intensité du gradient ϕ et la direction d_V de la chevelure et le sketch information de haut niveau représentant les particularités de la chevelure telle que les discontinuités.

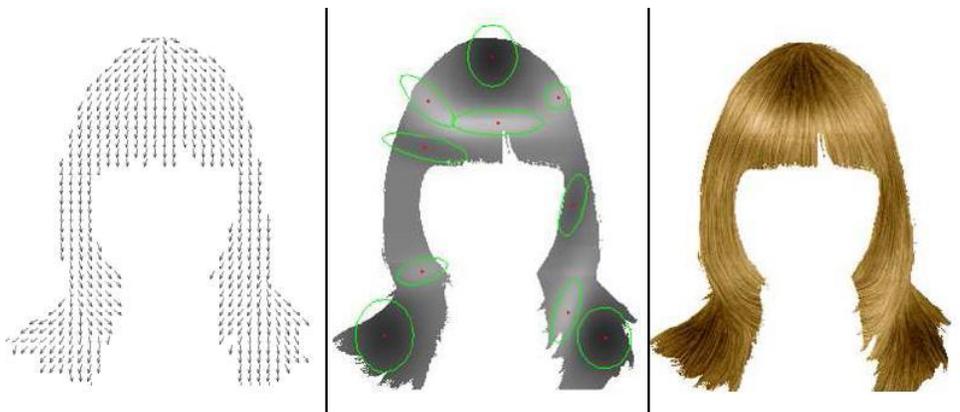


FIGURE 2.10 – **A Generative Sketch Model for Human Hair Analysis and Synthesis** : Exemple de coiffure synthétisée (c) à partir de la décomposition de 2 différentes cartes de caractérisation d'une chevelure : (a) la carte de direction, (b) la carte d'illumination ainsi que d'un modèle couleur (blond dans l'exemple). [Chen 06]

Ces travaux mettent en avant les caractéristiques de la chevelure pour la rendre réaliste : la couleur, l'illumination et la texture. La couleur étant déjà un paramètre largement utilisé pour la segmentation capillaire, son utilité pour décrire une chevelure n'est plus à démontrer. L'illumination peut être rattachée à l'analyse couleur pour la caractérisation de la chevelure, son étude pourrait permettre une meilleure définition des modèles couleurs pour mieux prendre en compte les variations de teinte au sein d'une même chevelure. La texture est aussi un élément caractéristique de la coiffure. Déjà à partir des travaux de Yacoob *et al* [Yacoob 06], nous avons émis l'hypothèse que la texture serait porteuse d'informations caractéristiques. Pourtant, elle n'est pour l'instant jamais prise en compte pour une application de segmentation capillaire.

2.1.7 Discussion Générale sur la segmentation capillaire

Nous venons de présenter l'ensemble des publications traitant de la segmentation de la chevelure que nous avons trouvé. Bien que la segmentation capillaire ne soit qu'à ses débuts nous

pouvons déjà dégager 3 caractéristiques principales de la littérature actuelle : la couleur, la localisation de la chevelure et la texture.

- **La couleur** : présente dans l'ensemble des techniques décrites, elle est la caractéristique de base de la segmentation capillaire. Son analyse permet la définition d'un classifieur couleur généralement modélisé par un mélange de gaussiennes [Zhang 08] [Lee 08] [Wang 09]. Le modèle couleur peut être calculé à partir de l'étude de fenêtres échantillons ou il peut être obtenu par apprentissage à partir de bases de vérités terrains segmentées manuellement [Zhang 08]. Une troisième approche est de combiner un modèle couleur initialisé par apprentissage puis de le mettre à jour pour l'adapter à l'image à segmenter [Lee 08] [Wang 09].
- **La localisation** : la localisation de la chevelure peut servir pour la définition de fenêtres échantillons afin de calculer les caractéristiques de la chevelure [Yacoob 06] [Rousset 08] où elle peut être utilisée comme paramètre à part entière de segmentation [Lee 08] [Wang 09]. Elle est particulièrement utile pour contrôler une segmentation basée uniquement sur l'analyse couleur. On peut localiser la chevelure soit par analyse géométrique à partir de données généralement calculées à partir de l'analyse du visage soit par apprentissage à partir d'une base de modèles de chevelure segmentés manuellement.
- **La texture** : elle est surtout utilisée pour la caractérisation de la chevelure [Yacoob 06] [Chen 06]. Mais c'est un paramètre de caractérisation qui porte beaucoup d'information qu'il pourrait être utile d'inclure dans un algorithme de segmentation.

Le choix de ces 3 paramètres peut être validé par une analyse de la perception naturelle d'une chevelure par un être humain. Il existe de nombreux adjectifs pour décrire une personne et le vocabulaire pour décrire une chevelure est aussi très riche. Il est possible de classer l'ensemble de ces adjectifs selon nos 3 critères :



FIGURE 2.11 – **Description d'une chevelure** : Description subjective d'une chevelure par l'utilisation de 3 adjectifs perceptifs caractéristiques (a) Coiffure aux cheveux courts, bruns et crépus (b) Coiffure carré long aux cheveux blonds et bouclés, (c) Coiffure aux cheveux longs, lisses et roux

- **La couleur** : Correspond aux adjectifs de couleur que l'on utilise pour décrire la teinte de la chevelure. Il existe une très grande variété de terme pour décrire l'ensemble des variations qu'une chevelure peut prendre : Aile de corbeau, Auburn, Blanc, Blond, Blond

vénitien, Brun, Châtain, Noir, Roux..

- **La localisation** : Correspond aux adjectifs décrivant le type de coupe de cheveux : long, court, mi-long, coupe au carré, rasé. Nous pouvons aussi inclure les coupes telles que les nattes, les couettes ou les queues de cheval.
- **La texture** : Correspond aux adjectifs décrivant plutôt la forme des cheveux plutôt que l'ensemble de la chevelure : crépu, bouclé, lisse, hirsute...

Nous avons défini comment caractériser la chevelure pour la localiser mais il faut aussi s'intéresser aux techniques de segmentation d'une image afin de pouvoir définir la meilleure approche permettant la meilleure définition d'un masque capillaire. La plupart des techniques utilisées sont orientées vers une segmentation par approche "régions" soit par une croissance itérative par blocs [Yacoob 06], soit par une diffusion de marqueurs [Rousset 08] [Wang 09] soit par la minimisation d'une fonction d'énergie [Lee 08].

Le choix de l'algorithme de segmentation étant aussi important que son initialisation et son contrôle par les paramètres capillaires, nous présenterons ici une revue non exhaustive des principales techniques de segmentation.

2.2 Techniques de Segmentation

La segmentation est une étape prépondérante dans le processus de traitement d'images. Elle vise à établir un partitionnement de l'image en régions ou en contours. Ces deux approches "*contours*" et "*régions*" sont duales en ce sens que les frontières d'une région définissent un contour et qu'un contour fermé définit une région. Elles amènent cependant à des algorithmes différents qui au final ne conduisent pas forcément aux mêmes résultats.

Avant la segmentation proprement dite, l'image est analysée lors d'une phase de pré-traitements pour déterminer des attributs ou des indices visuels de bas niveaux (par exemple des informations de gradient, de texture, de mouvement, de couleur...). Ces indices vont permettre d'orienter la segmentation pour des régions ayant le même type d'informations.

2.2.1 L'approche "Contours"

Les méthodes de segmentation par approche "contours" consistent à rechercher les discontinuités (variations d'intensité représentant des changements de propriétés physiques ou géométriques dans l'image) : variations d'illumination, ombres, changements d'orientation, variation de l'absorption de la lumière. En général ces variations d'intensité représentent des informations importantes qui permettent de déterminer les frontières des régions.

2.2.1.1 Les opérateurs dérivatifs

L'approche classique de détection de contours est l'approche dérivative. Cette approche est basée sur l'utilisation de dérivées de fonctions des niveaux de gris. Elle consiste à balayer l'image avec une fenêtre glissante définissant des opérateurs dérivatifs. À chaque position, l'opérateur est appliqué sur les pixels de la fenêtre afin de rechercher soit les extremums d'une dérivée première pour les opérateurs simples de type Prewitt ou Sobel, soit le passage à zéro d'une

dérivée seconde pour l'opérateur Laplacien. Un autre type d'approche plus récente repose sur la définition de critères d'optimalité de la détection de contours. Ces critères débouchent sur des filtres de lissage optimaux. On peut citer les opérateurs de Canny, de Shen-Castan et de Deriche qui sont basés notamment sur des calculs différentiels.

2.2.1.2 Les contours actifs

Nous pouvons définir aussi une approche de segmentation par contours actifs. Les contours actifs appelés également *snakes* sont une technique plus récente développée à la base par Kass *et al* en 1987. Cette méthode consiste à déterminer avec précision le contour d'un objet particulier dans une image.

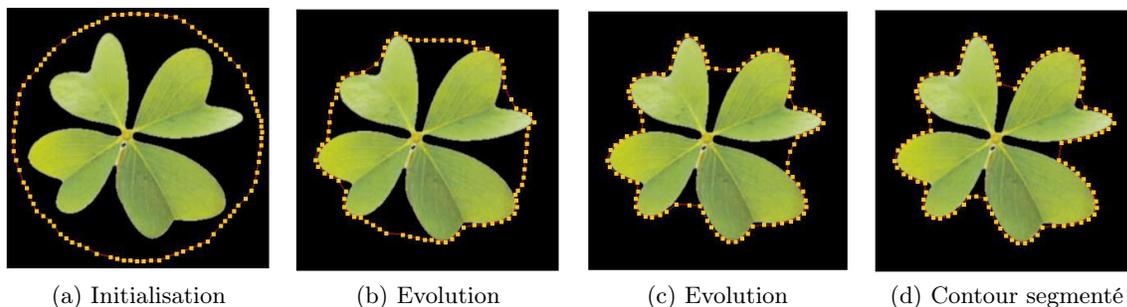


FIGURE 2.12 – **Contour actif** : Représentation du traitement de segmentation par contour actif : Pour la phase d'initialisation un contour grossier est défini (a), durant la phase d'évolution (b) et (c) le contour est déformé pour épouser la forme jusqu'à tendre à sa forme définitive (d) du contour finale.

Un contour actif est un ensemble de points qu'on va tenter de déplacer pour leur faire épouser une forme. L'idée de cette méthode est de déplacer les points pour les rapprocher des zones de forts gradients tout en conservant des caractéristiques comme la courbure du contour ou la répartition des points sur le contour ou d'autres contraintes liées à la disposition des points.

Au démarrage de l'algorithme, le contour est disposé uniformément autour de l'objet à détecter puis il va se rétracter pour en épouser au mieux ses formes. De la même manière, un contour actif peut aussi se dilater et tenter de remplir une forme, il sera alors situé à l'intérieur de celle-ci au démarrage de l'algorithme.

A chaque itération, l'algorithme va tenter de trouver un meilleur positionnement pour le contour pour minimiser les dérives par rapport aux contraintes utilisées. L'algorithme s'arrêtera lorsqu'il ne sera plus possible d'améliorer le positionnement ou simplement quand le nombre maximum d'itérations aura été atteint. On utilise les notions d'énergie interne et externe pour caractériser respectivement la forme du contour et tous les éléments qui lui sont propres ainsi que le positionnement du contour sur l'image en tenant compte des lignes de gradient.

2.2.1.3 Analyse des approches "contours"

Une rapide analyse de contours par les approches dérivatives permet une première caractérisation de la texture de la chevelure.

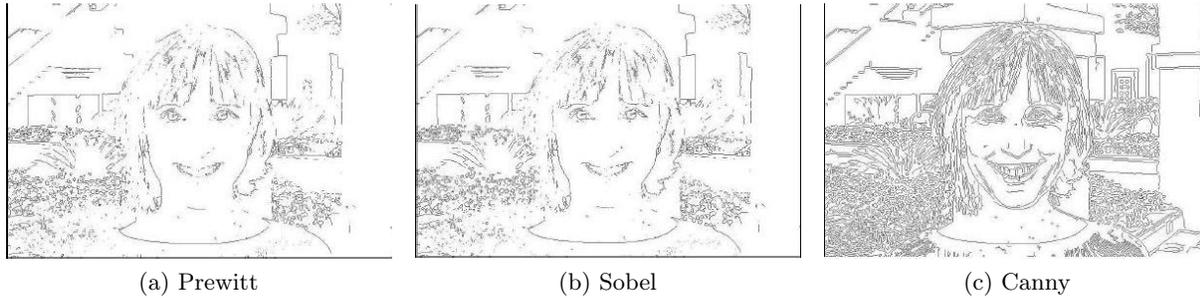


FIGURE 2.13 – "**Approche Derivative**" : Représentation de la détection de contours par les opérateurs dérivatifs de Prewitt (a), Sobel (b) et Canny (c)

Nous observons que les filtres figés par les opérateurs de Prewitt (*figure 2.13a*) et de Sobel (*figure 2.13b*) ne donnent pas une bonne détection des contours avec de nombreuses frontières perdues. Inversement le filtrage adapté par l'opérateur de Canny est très sensible et détecte un grand nombre de contours. Nous observons en particulier sur l'image obtenue par l'opérateur de Canny (*figure 2.13c*) que la chevelure est très texturée.

La difficulté à définir un gradient par les opérateurs dérivatifs, nous permet aussi de conclure sur l'approche par contours actifs. Dans un premier temps les équations d'évolution sont basées sur la maximisation d'une énergie le long d'un gradient et vu les résultats de ces opérateurs, deux cas de figure apparaissent :

- l'information gradient n'est pas assez importante et laisse des discontinuités dans les contours de la chevelure.
- il y a trop d'information et la chevelure est trop texturée.

Dans les deux cas, la sélection d'un contour optimal est difficile. De plus une des principales difficultés du traitement par contours actifs est l'initialisation du contour. En effet l'information utilisée pour le déplacement du contour étant de caractère local, une initialisation très proche de l'objet à segmenter est nécessaire. Initialisation particulièrement difficile car notre traitement doit pouvoir prendre en compte les nombreuses configurations de la chevelure sans intervention d'un opérateur pour le placement du contour initial.

La création d'un ensemble de modèles types de chevelure serait long et ne serait pas forcément adaptée à tous les cas de figure à cause du nombre important et presque infini des représentations possibles de la chevelure. Un autre problème posé par la segmentation par contours actifs est le nombre important d'itérations nécessaire pour avoir une convergence correcte du contour. Ce nombre d'itérations est d'autant plus important que l'objet à segmenter est irrégulier. Ce qui est particulièrement le cas pour la segmentation de chevelure.

Ainsi une segmentation par approche "contour" apparaît rapidement inadaptée pour la segmentation capillaire. Les techniques par approches "régions", utilisées dans la littérature, semblent plus appropriées.

2.2.2 Approche Régions

Les méthodes de segmentation par approche "régions" sont nombreuses. Elles consistent à rechercher des ensembles de pixels qui possèdent des propriétés communes. Les parties connexes de ces ensembles constituent les régions. Ces approches sont souvent réparties en 3 catégories : les méthodes par agrégation, les méthodes de division-fusion et les méthodes de modélisation statistique.

2.2.2.1 Les méthodes par agrégation

Elles consistent à élaborer les régions à partir du pixel en se basant sur des critères d'homogénéité. L'image est initialement décomposée en primitives 'régions' (une région pouvant être composée d'un seul pixel). Celles-ci sont ensuite regroupées de manière itérative selon un ou plusieurs critères de similarité jusqu'à la prise en compte de l'ensemble des pixels de l'image. On peut notamment citer l'algorithme de ligne de partage des eaux et les segmentations par similarité couleur mise en place par Yaccob *et al* [Yacoob 06] et par Wang *et al* [Wang 09].

Ligne de Partage des Eaux : LPE

Une image en niveaux de gris peut être considérée comme un relief accidenté où la valeur en niveaux de gris d'un pixel correspond à une altitude (sombre pour les basses altitudes, claires pour les hautes). Nous pouvons alors modéliser le relief en deux dimensions par des courbes de niveaux [Beucher 79] [Beucher 92] [Vincent 91].

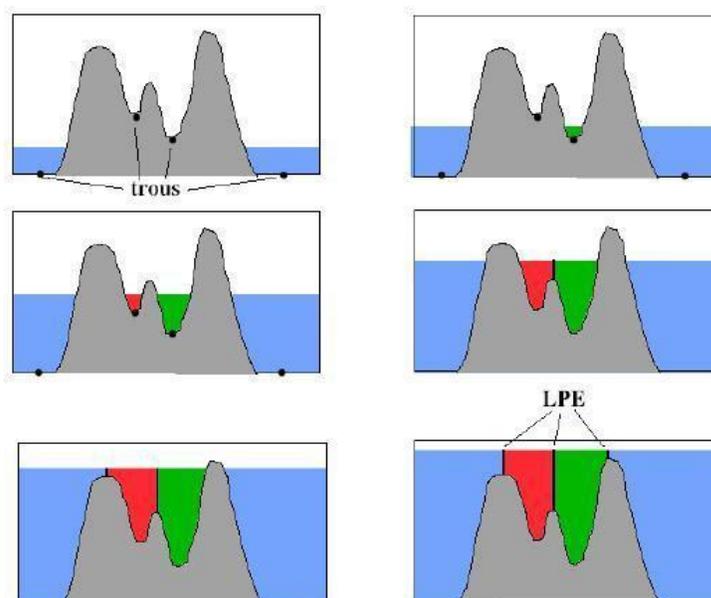


FIGURE 2.14 – **Ligne de partage des eaux LPE** : Schéma de l'évolution de l'algorithme de LPE : il est initialisé par la définition des trous pour les minimums locaux. Durant la phase d'évolution, le niveau de l'eau augmente. Des barrages sont créés afin d'éviter le mélange des eaux des différents bassins. La segmentation est obtenue, une fois que le niveau de l'eau a atteint son maximum, en projetant les contours des barrages.

Pour connaître les points les plus élevés, nous recherchons les lignes de crêtes qui peuvent

être trouvées par le calcul des maximums locaux du module des gradients dans la direction de celui-ci. Pour connaître les points les plus bas, nous recherchons les minimums locaux. Ces points correspondent aux lieux les plus bas atteignables sans être obligé de remonter. À chaque minimum local, nous associons un bassin versant qui correspond à la zone d'influence du minimum local. Toute goutte d'eau s'écoulant dans un bassin versant termine obligatoirement sa course au minimum local.

Le principe du traitement des lignes de partage des eaux est de déterminer les limites des zones d'influences de chaque minimum local en créant des lignes séparant les bassins versants les uns des autres. Une goutte positionnée sur une ligne de partage des eaux pourra, et seulement dans ce cas là, s'écouler vers deux bassins versants différents. Ce faisant, l'eau montante forme des lacs qui vont au fur et à mesure pouvoir fusionner entre eux. Afin d'éviter le mélange des eaux des différents bassins, nous construisons un barrage au point de fusion. Le phénomène de submersion se termine une fois l'ensemble de l'image sous l'eau. Pour récupérer les lignes de partage des eaux, nous projetons verticalement sur l'image l'ensemble des barrages. La projection des lacs forme les bassins versants.

2.2.2.2 Les méthodes de division-fusion

Ces méthodes sont aussi connues sous l'anglicisme '*split and merge*'. Après une étape d'initialisation où l'image est pré-segmentée grossièrement, le processus de segmentation itératif se divise en deux phases : une phase de division pour séparer toutes les régions non homogènes puis une phase de fusion de toutes les régions adjacentes dont l'union respecte des critères d'homogénéité

2.2.2.3 Les méthodes de modélisation statistique

Elles reposent sur la minimisation d'une fonction de vraisemblance (ou énergie). Cette fonction prend simultanément en compte la vraisemblance de l'appartenance du pixel à une région considérant son niveau de gris et les régions auxquelles appartiennent les pixels voisins. L'approche par Graph Cut utilisé par Lee *et al* [Lee 08] rentre dans cette catégorie.

2.2.2.4 Analyse des approches "régions"

Une méthode de segmentation par approche région semble plus appropriée pour la segmentation capillaire. D'ailleurs l'ensemble des travaux étudiés se basent sur ce type de techniques de segmentations [Yacoob 06] [Lee 08] [Wang 09]. Le principe généralement retenu est d'identifier des pixels cheveux comme germes potentiels du processus de segmentation et de diffuser ces germes en s'appuyant sur des fonctions de similarité ou d'énergie.

À partir de notre étude sur la segmentation par approche "contours", nous avons aussi exploré la piste d'une méthode par approche "contours" et "régions" combinées. Nous avons appliqué l'algorithme de la LPE pour des gradients de différentes tailles qui définissent la carte des lignes de crêtes. Les cartes de gradients étant obtenues par filtrage adapté par l'opérateur de Canny pour plusieurs résolutions, ce qui permettait de travailler sur plusieurs plages de fréquence (*figure 2.15*).

Nous avons rapidement abandonné cette approche car la présence de nombreux minimums locaux non significatifs entraînait une sur-segmentation de l'image, bien qu'en fonction de la résolution du filtre de Canny utilisé, nous pouvions contrôler la taille des bassins versants. Puis en se basant sur les travaux de Wang *et al* [Wang 09], nous aurions pu envisager l'utilisation de

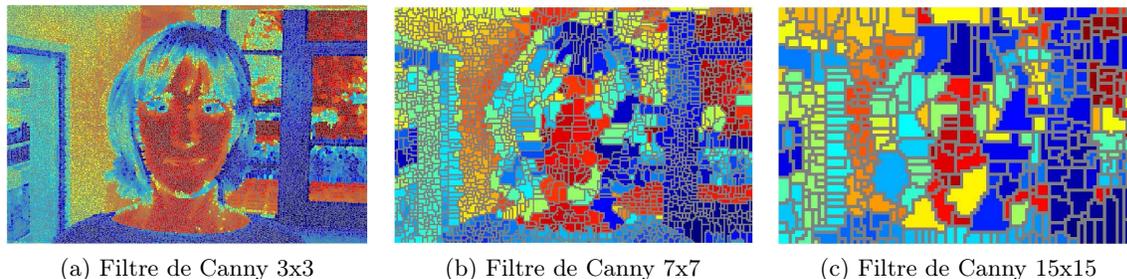


FIGURE 2.15 – **Segmentation par LPE** : Résultats de segmentation par lignes de partage des eaux initialisés à partir de cartes de gradients obtenues pour différentes résolutions du filtre de Canny

la LPE comme pré traitement afin d'obtenir une première segmentation grossière de la chevelure en ne conservant que les blocs qui pourraient correspondre à la chevelure. Mais il apparaît que les gradients les plus forts ne correspondent pas forcément aux contours de la chevelure et les segmentations obtenues fournissent des découpages de l'image incorrects. Une approche pour améliorer l'approche des lignes de partage des eaux consiste à supprimer certains minimums locaux en les remplaçant par des minimums régionaux. Ces minimums régionaux sont calculés par érosion géodésique qui consiste en un lissage local de l'image à partir de marqueurs définis à partir de l'étude des paramètres de la chevelure.

Le choix de notre algorithme de segmentation pour la segmentation capillaire s'est appuyé sur deux paramètres :

- Un algorithme qui permet la segmentation de l'image en deux classes afin d'obtenir une labellisation des pixels de l'image en 2 classes : la classe *Cheveux* H et la classe *Non cheveux* \bar{H} .
- Comme pour les autres approches de segmentation capillaire, nous initialiserons notre algorithme par l'identification de zones de chevelure pour le positionnement de germes.

Au cours de nos recherches de méthodes de segmentation par approches "régions", une technique a particulièrement retenu notre attention : la segmentation par traitement "*Matting*". Elle permet d'obtenir des segmentations précises de chevelures à partir d'un faible nombre de marqueurs représentant des germes qui seront diffusés dans l'image durant le traitement [Levin 06] [Levin 08].

2.3 Conclusion

Dans ce chapitre, nous avons présenté un état de l'art des travaux d'analyse et de segmentation capillaire. Puis nous avons introduit une revue non exhaustive des techniques de segmentation par approches "*contours*" et par approches "*régions*". Cette étude bibliographique a permis de cadrer le sujet et de définir les 2 grands axes de nos travaux pour la segmentation capillaire :

- Dans un premier temps, nous avons ainsi pu identifier 3 paramètres caractéristiques de la chevelure : la couleur, la texture et la localisation de la chevelure autour du visage. Ces 3 paramètres représentent des caractéristiques propres à la chevelure que nous allons

employer dans nos travaux. Dans le chapitre 4 nous présenterons notre analyse de texture de la chevelure, dans le chapitre 5 notre étude couleur et dans le chapitre 6 l'intégration de l'information de localisation de la chevelure dans notre processus de fusion de données.

- Dans un second temps nous avons pu choisir notre approche de segmentation par approche "régions" par l'utilisation de l'algorithme de "*Matting*" initialisé par les 3 paramètres de la chevelure que nous avons définis précédemment.

3 Cadre de la Segmentation Capillaire



Sommaire

3.1	Digital Matting ou Composition d'images	43
3.1.1	État de l'art sur le " <i>Matting</i> "	44
3.1.1.1	Approche " <i>Matting</i> " par fond de couleur constante	44
3.1.1.2	La Trimap	45
3.1.1.3	Approche Knock Out	46
3.1.1.4	Approche Bayésienne	47
3.1.1.5	Approche par méthode Poisson	47
3.1.1.6	Approche par Flash " <i>Matting</i> "	47
3.1.1.7	Approche " <i>Matting</i> " par marqueurs	48
3.1.2	Bilan Matting	48
3.2	"<i>Matting</i>" par marqueur	49
3.2.1	Images en niveau de gris	49
3.2.2	Images couleur	50
3.3	Etude du "<i>Matting</i>" par marqueur	52
3.3.1	Image Synthétique	52
3.3.2	Images Naturelles	55
3.4	Cadre d'étude	57
3.4.1	Formalisme	57
3.4.2	Hypothèse de travail	58
3.4.3	Choix des bases d'images	59
3.5	Évaluation des résultats	62
3.5.1	Base de Vérités Terrains	62
3.5.2	Évaluation de la segmentation capillaire	64
3.6	Conclusion	64

Nous avons conclu le chapitre précédent en choisissant pour notre application de segmentation capillaire un algorithme de segmentation par croissance de "régions" appelé traitement "*Matting*".

Dans ce chapitre, nous commencerons par une présentation plus approfondie du traitement "*Matting*" (*section 3.1*). Dans un second temps, nous introduirons en détail la méthode développée par Levin *et al* (*section 3.2*). Puis nous testerons les avantages et inconvénients de cette approche afin de définir les limites de l'approche "*Matting*" (*section 3.3*), ce qui nous permettra de définir un cadre de travail pour notre étude de segmentation capillaire (*section 3.4*). Enfin nous présenterons notre méthode d'évaluation de résultat ainsi que la méthode d'acquisition de notre base de référence (*section 3.5*).

3.1 Digital Matting ou Composition d'images

La composition d'images numériques ou "*Matting*" consiste au "mélange" de plusieurs images en une unique image composite C finale. C'est l'évolution technologique issue de la composition d'images optiques.

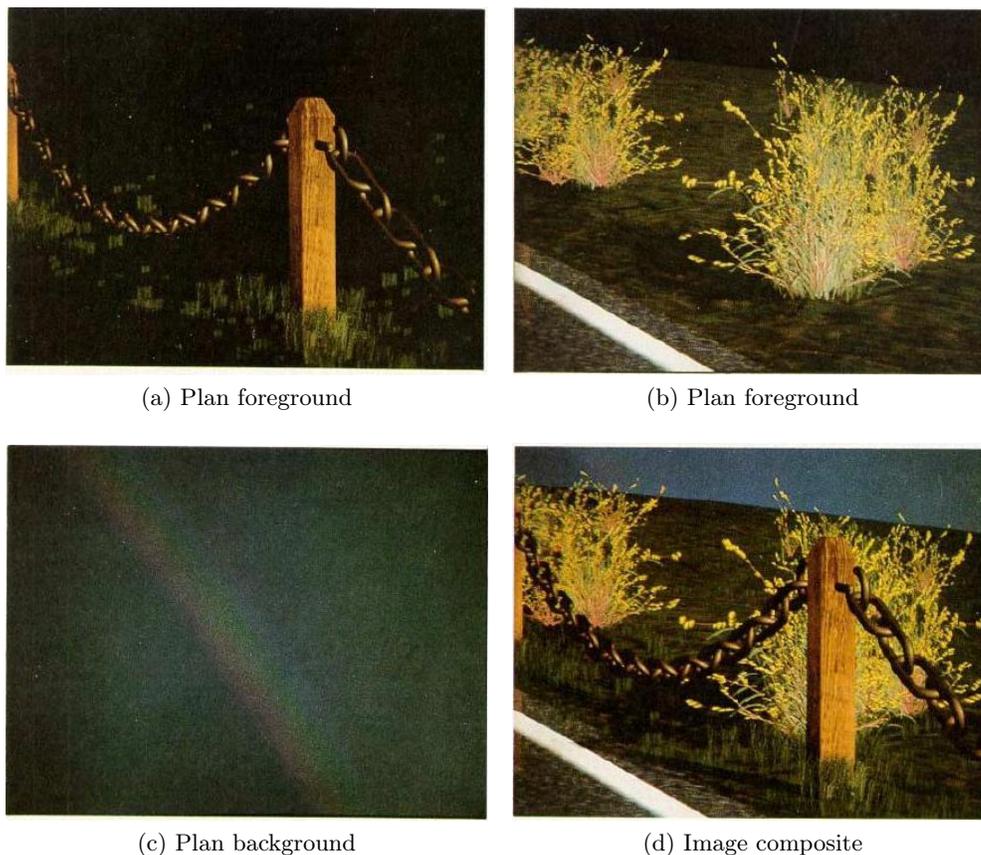


FIGURE 3.1 – "**Compositing Digital Image**" : Exemple de composition d'image (d) à partir de 2 images d'information ou de premier plan F (a) et (b) et d'une image de fond B (c) [Porter 84]

Le processus de composition d'image est utilisé dans l'industrie filmographique depuis le dé-

but des années 1970. Ce procédé consiste, à partir d'une image, à segmenter la partie que l'on veut récupérer ou partie "d'information" et de retirer le reste de l'image pour pouvoir le remplacer par une autre image (figure (3.1)). Au début, ce procédé consistait à rajouter "un matte", une bande transparente à certains endroits de l'image. Ainsi en combinant cette bande avec la bande originale lors de la projection, la lumière passait à travers les zones transparentes et était bloquée ailleurs. On obtenait alors uniquement la projection des zones voulues.

Pour combiner ces différents plans images, il existe un grand nombre d'opérations possibles. Le procédé le plus utilisé est l'"over opérateur" qui est défini pour chaque pixel de l'image composite \mathbf{C} par l'équation suivante :

$$C = \alpha F + (1 - \alpha)B. \quad (3.1)$$

avec

- la zone d'information ou zone de premier plan *Foreground* \mathbf{F} :
- la zone de fond *Background* \mathbf{B} :
- le paramètre d'opacité α : En 1984, Porter et Duff [Porter 84] introduisent le paramètre α , la composante linéaire de mélange liant le premier plan et le fond, pour la création d'images synthétiques complexes. Ainsi la composante numérique correspondant à l'opacité était introduite. Un pixel avec une valeur α de 1 est opaque et inversement un pixel avec une valeur α de 0 est transparent.

À la base le traitement "Matting" servait à la composition d'image. Le mot "Matting" vient de l'anglais "to mat" qui se traduit en français par le verbe "emmêler". Ce sens peut s'expliquer par l'interprétation qu'une image composite \mathbf{C} est l'entremêlement de l'image \mathbf{F} de premier plan et de l'image \mathbf{B} du plan du fond. À partir de cette définition, la segmentation analogue revient alors à "démêler" l'image composite \mathbf{C} en estimant le paramètre α à chaque pixel puis de séparer les deux plans images \mathbf{F} et \mathbf{B} .

3.1.1 État de l'art sur le "Matting"

Historiquement le "Matting" est issu de l'industrie filmographique et la plupart des bases de ce procédé ont vu le jour bien avant l'avènement de l'informatique. Ainsi on peut attribuer à Vlahos [Berman 00] les bases de ce procédé. Personne d'expérience dans son milieu, il a notamment résolu les premiers problèmes liés à la saturation des blancs ou liés à la couleur de peau.

3.1.1.1 Approche "Matting" par fond de couleur constante

En 1996, Smith et Blinn [Smith 96] ont créé le formalisme mathématique afin de pouvoir mieux définir les travaux de Vlahos. Mais l'équation de composition (3.1) est un problème quadratique mal posé avec une équation pour trois inconnues.

Il faut alors la définition de connaissances pour la résolution du système. Soit par l'apport d'information sur les valeurs α ou sur les plans \mathbf{F} et \mathbf{B} . La solution la plus utilisée est l'utilisation d'un fond uniforme de couleur non naturelle saturée en particulier le bleu ou le vert pur de la décomposition RVB.

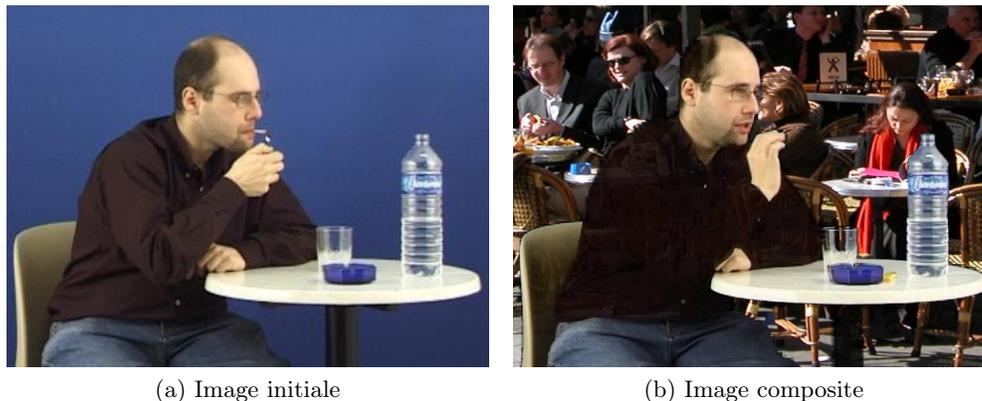


FIGURE 3.2 – "*Matting*" par fond de couleur constante : Exemple de composition d'image (b) utilisant un fond de couleur uniforme bleu (a)

L'équation (3.1) est simplifiée par la connaissance du fond \mathbf{B} . Le système devient alors résoluble et le problème revient à estimer le paramètre α pour l'ensemble des pixels afin de créer le plan de mélange appelé "*alpha matte*" entre le plan \mathbf{F} et le fond \mathbf{B} .

Son utilisation est connue du grand public car c'est le procédé utilisé généralement pour la création des effets spéciaux dans l'industrie cinématographique (*figure 3.2*).

3.1.1.2 La Trimap

Pour une image naturelle, aucun des paramètres de l'équation (3.1) ne sont connus. Afin de pouvoir résoudre le système, il faut rajouter de l'information soit par l'intervention d'un opérateur soit par de la connaissance a priori comme l'utilisation d'une *Trimap*.

La plupart des techniques récentes se basent sur l'utilisation d'une *Trimap* [Ruzon 00] [Chuang 01] [Karlsson 04] [Tang 06]. Cette *Trimap* est une segmentation grossière (typiquement faite manuellement) de l'image en 3 zones : la partie connue du premier plan (codée en blanc), la partie connue du fond (codée en noir) et la partie inconnue où l'alpha matte reste à être estimé en gris. Pour une *Trimap* donnée, ces méthodes résolvent les plans \mathbf{B} et \mathbf{F} et α simultanément en alternant l'estimation de \mathbf{F} et \mathbf{B} avec celle du plan α .

La *Trimap* a été introduite la première fois par Ruzon et Tomasi [Ruzon 00]. Dans un premier temps la région inconnue est partitionnée en sous régions basées sur des points d'ancrage (*figure 3.3a*). Chaque point d'ancrage correspond au centre d'une fenêtre qui définit une distribution locale de couleur. La taille des fenêtres est déterminée de façon à englober des régions de zones connues de fond \mathbf{B} et de premier plan \mathbf{F} . Les pixels sont séparés en groupes pour chaque zone connue et serviront à estimer leurs distributions, où chaque groupe est modélisé comme un mélange de gaussiennes alignées selon un axe (*figure 3.3d*). Le modèle couleur de l'image composite sur l'ensemble de la fenêtre est alors modélisé par une gaussienne alignée sur les 2 autres gaussiennes précédemment estimées. La distance entre chaque distribution correspond à la valeur en α du pixel. La valeur en α est choisie pour maximiser la probabilité du pixel étudié. Une fois la valeur de α estimé, les valeurs des plans \mathbf{F} et \mathbf{B} du pixel sont calculés comme la moyenne de chaque groupe de la fenêtre d'étude.

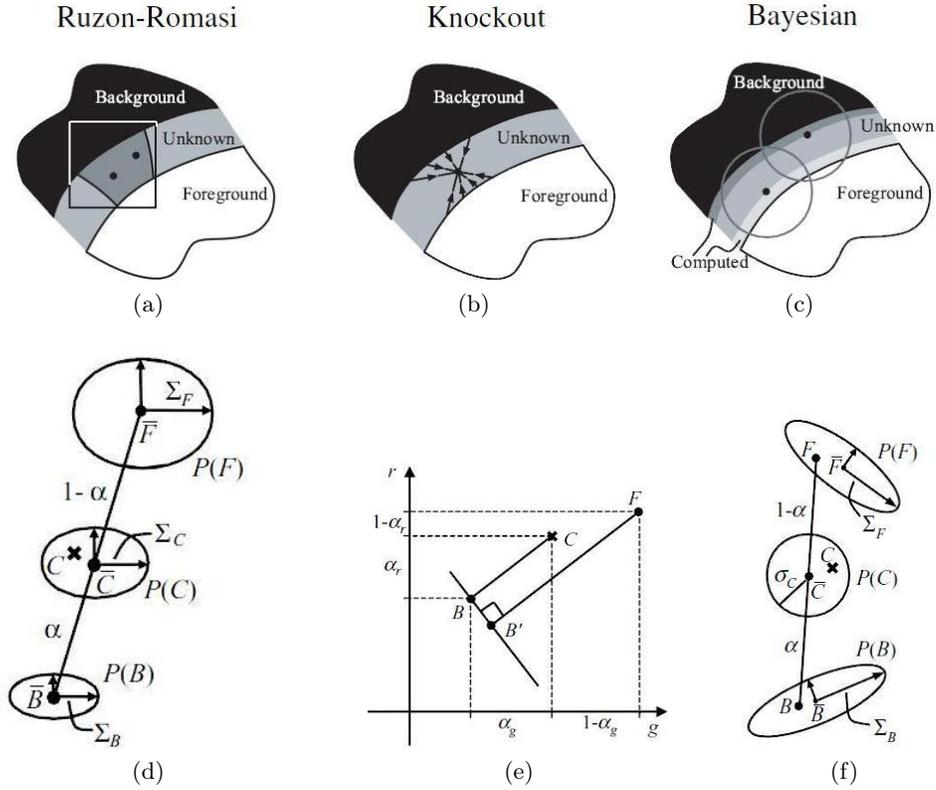


FIGURE 3.3 – **Trimap "Matting"** : (1^{ère} ligne) La *Trimap* est une segmentation grossière (typiquement faite manuellement) de l'image en 3 zones : la partie connue du premier plan (codée en blanc), la partie connue du fond (codée en noir) et la partie inconnue où l'alpha matte reste à être estimé en gris. (2^{ème} ligne) Schéma de la modélisation des distributions des plans \mathbf{F} et \mathbf{B} pour l'estimation du plan α . [Chuang 01]

3.1.1.3 Approche Knock Out

Pour la méthode Knockout, les plans \mathbf{F} et \mathbf{B} sont supposés uniformes et les pixels de la zone inconnue sont extrapolés à partir des pixels connus des plans \mathbf{F} et \mathbf{B} (figure 3.3b). Par exemple pour un pixel de la région inconnue, le plan \mathbf{F} (respectivement \mathbf{B}') est calculé comme la somme pondéré des pixels contours de la zone connue de \mathbf{F} en pondérant le pixel connu le plus proche à 1 et décroissant linéairement jusqu'à 0 pour les pixels se trouvant au double de la distance du plus proche pixel. On redéfinit ensuite la valeur de \mathbf{B} en projetant parallèlement au model line color \mathbf{FB}' la valeur du pixel inconnue sur un plan normal en \mathbf{B}' (figure 3.3e). Finalement l'alpha matte est estimé suivant la relation suivante :

$$\alpha_{r,g,b} = \frac{f(C_{r,g,b}) - f(B_{r,g,b})}{f(F_{r,g,b}) - f(B_{r,g,b})} \quad (3.2)$$

où $f()$ est la fonction projetant respectivement chaque composante RGB sur un axe. L'équation (3.2) est obtenu facilement en développant l'équation de composition (3.1)

$$C = \alpha F + (1 - \alpha) B \iff C - B = \alpha F - \alpha B \iff \alpha = (C - B)/(F - B) \quad (3.3)$$

3.1.1.4 Approche Bayésienne

Le "*Matting*" par approche bayésienne [Chuang 01] [Karlsson 04] est une amélioration des travaux de Ruzon et Tomasi [Ruzon 00]. La méthode diffère peu mais améliore néanmoins fortement les résultats. Premièrement les distributions couleurs des pixels connus de \mathbf{F} et \mathbf{B} sont modélisées par des gaussiennes orientées (*figure 3.3f*). Deuxièmement l'approche utilisée pour définir ces gaussiennes est basée sur des fenêtres glissantes à partir des régions connues vers le centre de la zone inconnue (*figure 3.3c*). Ainsi les distributions gaussiennes orientées sont calculées à partir des pixels des régions connues (comme pour Ruzon et Tomasi) mais aussi à partir des valeurs déjà calculées. Enfin l'estimation des valeurs de α , \mathbf{F} et \mathbf{B} est effectuée simultanément plutôt qu'en alternance.

Finalement l'approche "*Matting*" Bayésienne se base sur une résolution par une approche MAP (Maximum A Posteriori) qui revient à maximiser l'équation suivante :

$$\arg \max_{F,B,\alpha} P(F, B, \alpha | C) = \arg \max_{F,B,\alpha} \frac{P(C|F, B, \alpha)P(F)P(B)P(\alpha)}{P(C)} \quad (3.4)$$

3.1.1.5 Approche par méthode Poisson

L'approche "*Matting*" par la méthode Poisson [Tang 06] utilise aussi une Trimap comme carte d'initialisation. L'estimation de l'alpha matte est basée sur la résolution un système d'équation de Poisson grâce au champ de gradient de l'alpha matte et des conditions aux frontières de Dirichlet, le champ de gradient de l'alpha matte étant approximé à partir du gradient de l'équation de composition. Puis l'alpha mate est estimé en résolvant l'équation dont les gradients approximent le mieux possible le champ de gradient calculé.

3.1.1.6 Approche par Flash "*Matting*"

Une autre approche développée par Sun *et al* [Sun 06] [Sun 07] est basée sur une observation : si le fond d'une image est assez éloigné, les effets d'un flash ne sont visibles que sur les éléments du premier plan. Ainsi à partir de 2 images : l'image ambiante sans flash et l'image avec flash, on peut définir un nouveau système d'équation :

$$C = \alpha F + (1 - \alpha)B. \quad (3.5)$$

$$C^f = \alpha F^f + (1 - \alpha)B^f. \quad (3.6)$$

avec F et B les plans pour l'image ambiante et F^f et B^f pour l'image avec flash. L'hypothèse est faite que la prise de vue reste identique entre les 2 photos et donc que les valeurs du plan α restent inchangées.

On considère l'équation exprimant la radiance E de l'illumination par un flash d'intensité lumineuse L sur un point d'une surface P :

$$E = L \cdot \phi(\omega_i, \omega_o) \cdot r^{-2} \cdot \cos\theta \quad (3.7)$$

avec $\phi(\omega_i, \omega_o)$ la fonction de reflectance bidirectionnelle d'une surface, r la distance au flash et θ l'angle entre la direction du flash et la surface P .

La loi en inverse au carré de r montre que l'intensité lumineuse du flash décroît fortement avec la distance r . Ainsi si la distance entre le flash et le fond de l'image est grande, la différence

d'illumination entre les 2 images avec ou sans flash sera minimale et nous pourrions considérer alors $B \approx B^f$. L'équation (3.6) peut alors être redéfinie par l'équation :

$$C^f = \alpha F^f + (1 - \alpha)B \quad (3.8)$$

En soustrayant l'équation (3.5) de l'équation (3.8), on définit l'équation dite "*Foreground Flash Matting*" :

$$C' = C^f - C = \alpha(F^f - F) = \alpha F' \quad (3.9)$$

Au final, chaque pixel est estimé en maximisant la fonction de Log vraisemblance $L(\alpha, F, B, F'|C, C')$.

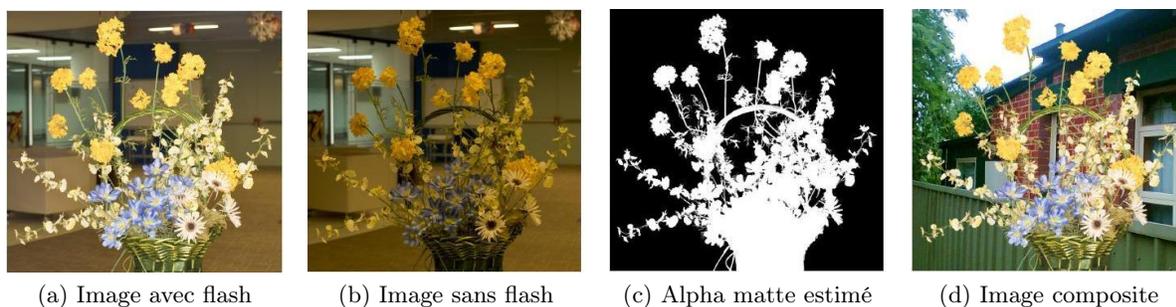


FIGURE 3.4 – **Flash Matting** : Exemple de composition (d) par Flash Matting en utilisant l'alpha matte estimé (c) à partir de la comparaison entre la même image prise avec (a) et sans (b) flash [Sun 06].

3.1.1.7 Approche "*Matting*" par marqueurs

Pour finir nous présentons la méthode développée par Levin *et al* [Levin 06] [Levin 08] basée sur leurs travaux de colorisation [Levin 04]. Afin de résoudre le problème d'évaluation de la *Trimap*, ils ont développé une approche ne nécessitant qu'un faible nombre de pixels identifiés : les marqueurs. Il suffira de marquer sommairement les zones d'information \mathbf{F} et de fond \mathbf{B} connus et de diffuser ces marqueurs dans l'image suivant des conditions d'uniformité locale de la luminance ou de la couleur. Le traitement "*Matting*" par marqueurs sera décrit plus en détails dans la section (*section 3.2*).

3.1.2 Bilan Matting

La segmentation d'image par approche "*Matting*" est très performante. Dans la figure 3.5, nous pouvons observer des résultats de Levin *et al* issus de la comparaison avec l'approche par "*Matting*" Bayésien et par approche Poisson de leurs résultats de segmentation sur 2 exemples [Levin 06]. Nous observons des segmentations très performantes. La segmentation des poils de queue du chien du premier exemple et tout particulièrement des poils des oreilles du koala dans le deuxième exemple sont très précises. Ces résultats sont intéressants car ils montrent que le traitement "*Matting*" est adapté pour la segmentation capillaire et permettra une segmentation fine des mèches de cheveux.

Toutes ces méthodes donnent de bons résultats. Cependant, pour les méthodes utilisant une *Trimap*, les résultats dépendent directement de la définition de la zone inconnue. Plus cette zone

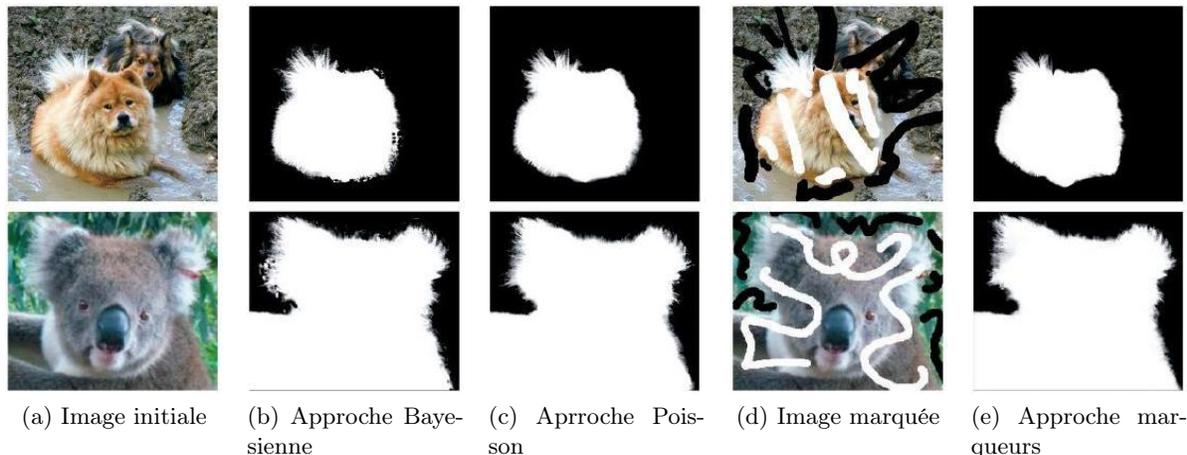


FIGURE 3.5 – **Traitements "Matting" - Comparaisons de résultats** : Exemple d'alpha matte obtenus par les approches de "Matting" Bayésien, par approche Poisson et par approche marqueurs [Levin 06].

sera petite, plus la segmentation sera performante. Il faut donc que la connaissance des zones **F** et **B** soient importantes. Dans notre approche de segmentation capillaire, la définition automatique d'une Trimap précise adaptée n'est pas envisageable car elle demanderait déjà une trop grande connaissance de la chevelure à segmenter. Définition d'autant plus difficile car il faudrait prendre en compte la grande diversité des coiffures possibles.

L'approche "Matting" par marqueurs apporte une alternative performante. Les segmentations obtenues sont aussi performantes. Mais la définition de l'information connue des plans **F** et **B** est facilitée par le marquage d'un faible nombre de pixels.

3.2 "Matting" par marqueur

Dans cette section nous présenterons en détails les travaux développés par Anat *et al* [Levin 06] [Levin 08]. Dans un premier temps, nous verrons l'approche pour une image en niveau de gris (*section 3.2.1*). Dans un second temps, cette solution sera étendue pour les modèles couleurs (*section 3.2.2*).

3.2.1 Images en niveau de gris

Pour pouvoir résoudre le système d'équation il faut certaines hypothèses sur la nature des plans **F**, **B** et α .

Les plans **B** et **F** sont considérés comme localement uniformes sur une petite fenêtre autour de chaque pixel. L'hypothèse faite que les plans **F** et **B** sont localement uniformes n'implique aucunement que l'image composite **C** le soit également. En effet les discontinuités présentes dans le plan α seront retrouvées pour le plan **C**.

On peut alors redéfinir l'expression de α par l'équation linéaire suivante :

$$\alpha_i = aC_i + b, \forall i \in \omega \quad (3.10)$$

avec $a = \frac{1}{F-B}$, $b = \frac{-B}{F-B}$ et ω une fenêtre.

L'estimation du traitement "Matting" revient alors à la recherche des paramètres α , a et b qui minimisent la fonction de coût :

$$J(\alpha, a, b) = \sum_{j \in I} \left(\sum_{i \in \omega_j} (\alpha_i - a_j C_i - b_j)^2 + \varepsilon a_j^2 \right) \quad (3.11)$$

avec ω_j une fenêtre centrée en j ¹.

Un terme de régulation en a a été rajouté afin de permettre une stabilité numérique pour les cas où la j^{ime} fenêtre serait constante. De plus, minimiser la norme de a biaise la solution vers un α plus lisse. En effet $a_j = 0$ implique que α soit constant pour toute la j^{ime} fenêtre. Pour la résolution de l'équation, des fenêtres 3x3 sont utilisées. Comme l'équation est résolue pour chaque pixel, on obtient un chevauchement des fenêtres ω_j . Ainsi grâce à cette propriété, l'information peut être propagée de pixel voisin en pixel voisin.

On peut ainsi éliminer les paramètres a et b et obtenir une équation quadratique en α avec seulement N inconnues :

$$J(\alpha) = \min_{a,b} J(\alpha, a, b) \quad (3.12)$$

Équation que l'on peut mettre de la forme :

$$J(\alpha) = \alpha^T L \alpha \quad (3.13)$$

Où L est une matrice $N \times N$ qui pour chaque indice (i, j) définit alors la valeur $L(i, j)$ selon l'équation :

$$L(i, j) = \sum_k \left(\delta_{ij} - \frac{1}{|\omega_k|} \left(1 + \frac{1}{\frac{\varepsilon}{|\omega_k|} + \sigma_k^2} (C_i - \mu_k)(C_j - \mu_k) \right) \right) \quad (3.14)$$

avec δ_{ij} le symbole kronecker, μ_k et σ_k respectivement la moyenne et la variance de la fenêtre ω_k autour du pixel k . et $|\omega_k|$ définit le nombre de pixels de la fenêtre (typiquement 9 pour des imagettes 3x3).

L'estimation du plan de l'alpha matte revient à la résolution d'une équation quadratique dépendant de la moyenne et variance pour une fenêtre.

3.2.2 Images couleur

On peut étendre le problème aux images couleurs où la définition de l'équation de composition (3.1) s'écrit alors sous la forme d'un système à 3 équations pour 7 inconnues

$$\begin{aligned} C_r &= \alpha F_r + (1 - \alpha) B_r \\ C_v &= \alpha F_v + (1 - \alpha) B_v \\ C_b &= \alpha F_b + (1 - \alpha) B_b \end{aligned} \quad (3.15)$$

Une approche simple serait alors d'appliquer la méthode en niveaux de gris à chaque composante indépendamment. Dans [Levin 06] [Levin 08], Levin *et al* remplacent ce modèle linéaire (3.10) par un modèle linéaire en 4 dimensions :

$$\alpha_i \approx \sum_c a^c C_i^c + b, \forall i \in \omega \quad (3.16)$$

1. On reconnait ici une équation de la forme de la méthode des moindres carrés.

avec c représentant la somme de tous les plans couleurs.

L'avantage de ce modèle linéaire est qu'il détend l'hypothèse d'uniformité locale des plans \mathbf{F} et \mathbf{B} . Pour une petite fenêtre d'étude, les plans \mathbf{F} et \mathbf{B} sont supposés être le mélange linéaire de seulement 2 couleurs. Ainsi pour une petite fenêtre, les valeurs de \mathbf{F} (respectivement de \mathbf{B}) apparaissent sous la forme d'une seule ligne dans l'espace RGB vérifiant l'équation $F_i = \beta_i^F F_1 + (1 - \beta_i^F) F_2$ (respectivement $B_i = \beta_i^B B_1 + (1 - \beta_i^B) B_2$) (figure 3.6).

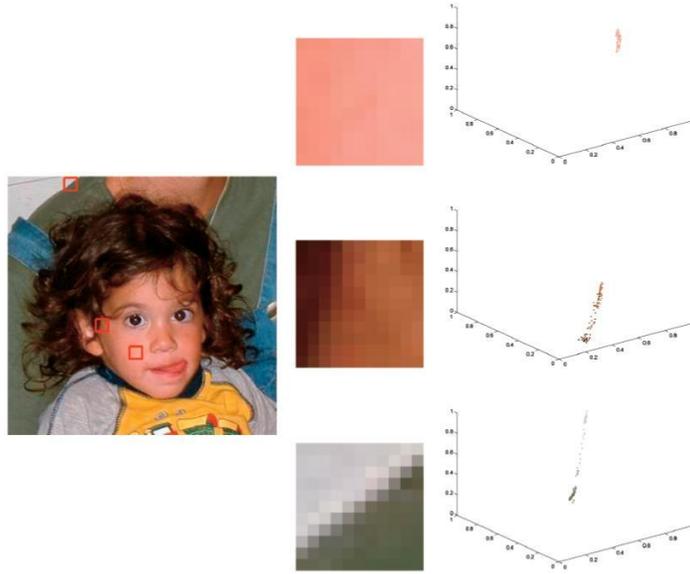


FIGURE 3.6 – **Uniformité locale de la distribution couleur** : Représentation de la distribution couleur dans le plan RVB pour une petite fenêtre échantillon [Levin 06].

Si le premier plan \mathbf{F} et le fond \mathbf{B} satisfont cette hypothèse pour une petite fenêtre alors on peut poser l'égalité dans l'équation (3.16)

$$\alpha_i = \sum_c a^c C_i^c + b, \forall i \in \omega \quad (3.17)$$

Avec cette nouvelle définition du modèle linéaire (3.17), la fonction de coût pour le "Matting" des images couleur s'exprime selon l'équation :

$$J(\alpha, a, b) = \sum_{j \in I} \left(\sum_{i \in \omega_j} (\alpha_i - \sum_c a_j^c C_i^c - b_j)^2 + \varepsilon \sum_c a_j^{c2} \right) \quad (3.18)$$

Comme pour le modèle des images en niveaux de gris, le système peut être simplifié et les termes a^c et b peuvent être retirés de la fonction de coût pour retrouver un problème quadratique modélisable selon l'équation (3.13) où la matrice s'exprime alors selon l'équation :

$$L(i, j) = \sum_{k | (i, j) \in \omega} \left(\delta_{ij} - \frac{1}{|\omega_k|} \left(1 + (C_i - \mu_k)(\Sigma_k + \frac{\varepsilon}{|\omega_k|} I_3)(C_j - \mu_k) \right) \right) \quad (3.19)$$

avec δ_{ij} le symbole Kronecker, μ_k , σ_k et Σ_k respectivement la moyenne, la variance et la matrice de covariance de la fenêtre ω_k autour du pixel k . $|\omega_k|$ définit le nombre de pixels de la fenêtre

(typiquement 9 pour des imagettes 3x3) et I_3 la matrice identité.

Après l'estimation de l'alpha matte, il est normalement nécessaire de calculer aussi le plan \mathbf{F} et parfois le plan \mathbf{B} . Une approche pour reconstruire les plans \mathbf{F} et \mathbf{B} est de résoudre l'équation(3.16) pour des valeurs de a et b optimales calculées en utilisant la méthode des moindres carrés. Pour des modèles complexes de premier plan \mathbf{F} et de fond \mathbf{B} , une telle reconstruction peut produire des résultats bruités, c'est pourquoi Levin *et al* résolvent l'estimation des plans \mathbf{F} et \mathbf{B} en conservant les hypothèses d'uniformité couleur des plans \mathbf{F} et \mathbf{B} . Comme pour l'étude de segmentation capillaire, nous utilisons l'alpha matte pour la définition de notre masque de segmentation. Le lecteur intéressé trouvera plus de détails de l'estimation des plans \mathbf{F} et \mathbf{B} dans [Levin 06] [Levin 08].

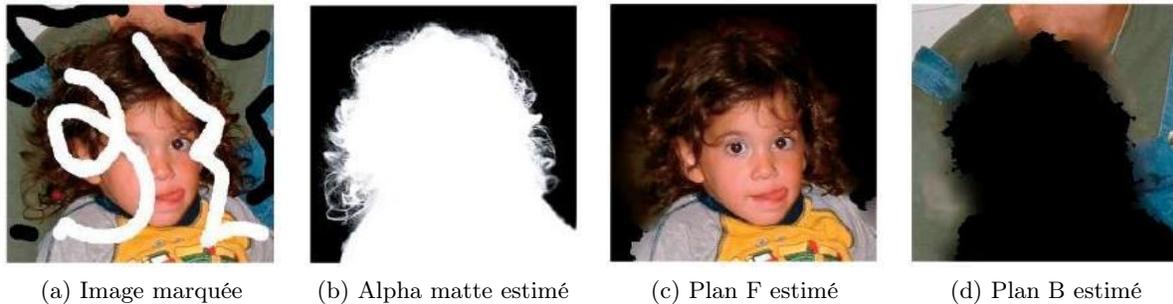


FIGURE 3.7 – **Approche "Matting" par marqueurs** : Résultats de l'estimation des plans α , \mathbf{F} et \mathbf{B} pour une image couleur [Levin 06].

3.3 Etude du "Matting" par marqueur

Nous avons présenté dans la section précédente la théorie du "Matting" par marqueurs qui permet avec un très faible nombre de pixels identifiés d'estimer un plan alpha matte qui segmente une image en 2 zones. En effet même si l'algorithme donne de bons résultats pour un faible nombre de marqueurs, l'interaction, le placement et le nombre de ces marqueurs doivent être pris en compte pour pouvoir initialiser le traitement au mieux grâce aux caractéristiques de la chevelure. Nous présentons ici plusieurs traitements pour évaluer les performances de l'algorithme et nous permettre de définir un cadre d'utilisation.

3.3.1 Image Synthétique

Dans un premier temps nous avons testé l'algorithme sur des images synthétiques afin de limiter l'étude au seul rôle des marqueurs dans la segmentation et non par l'information contenue dans l'image. Nous considérons une image couleur coupée en deux : avec la partie bleue considérée comme la zone de premier plan \mathbf{F} et la partie rouge considérée comme la zone de fond \mathbf{B} .

Nous allons ainsi tester l'influence du placement et du nombre de marqueurs. Nous testerons aussi la dynamique d'exécution de l'algorithme en le testant pour plusieurs résolutions d'images.

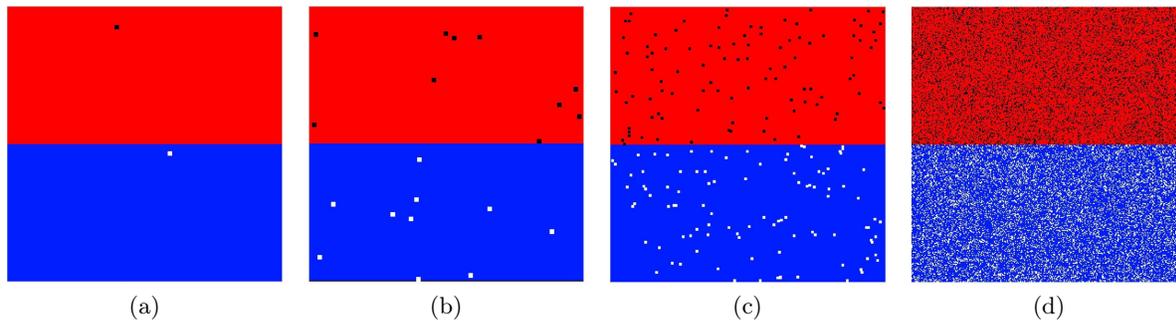


FIGURE 3.8 – Étude du "Matting" - Nombre de marqueurs : Exemples d'images synthétiques pour un nombre grandissant de pixels marqués.

- **Nombre de marqueurs** : le premier paramètre que nous avons testé est l'influence du nombre de pixels marqués dans chaque zone pour l'estimation de l'alpha matte. Cette étude est particulièrement importante car nous avons choisi l'approche "Matting" justement à cause de la possibilité d'avoir un nombre de points marqués limités. Cette étude va nous permettre de définir deux caractéristiques :
 - Le nombre minimal de pixels devant être marqués pour assurer une bonne segmentation.
 - L'influence du nombre de pixels sur la segmentation (figure 3.8). En effet même si un faible nombre de pixels est nécessaire pour obtenir une bonne segmentation, nous devons définir quelles sont les améliorations obtenues pour un grand nombre de pixels marqués.

Même pour des marqueurs définis par un seul pixel (figure 3.8a), nous obtenons une bonne segmentation de l'image synthétique.

Comme nous obtenons une segmentation réussie pour des marqueurs de seulement un pixel, nous ne pouvons pas tester l'influence du nombre de pixels représentant les marqueurs sur la qualité de la segmentation. Par contre, nous avons mesuré le temps de calcul de l'algorithme pour effectuer le traitement.

La figure 3.9 représente l'évolution du temps de calcul (en ordonnée) en fonction du pourcentage de pixels marqués (en abscisse) dans chacune des zones **F** et **B**. Pour une image de taille de 300×300 pixels, nous obtenons un temps de calcul d'environ 14.2 seconde pour des marqueurs représentés par un unique pixel. Ce temps de calcul est à peu près constant jusqu'à ce que les pixels marqués représentent environ 30% de chaque zone. Puis nous observons une légère augmentation du temps de calcul qui tends vers les 15 secondes pour des marqueurs représentant environ 75% des pixels. Enfin le temps de calcul décroît rapidement pour tendre vers un temps de calcul nul pour l'ensemble des pixels marqués. En effet si tous les pixels sont identifiés, l'estimation de l'alpha matte est alors immédiate.

- **Placement des marqueurs** Nous venons de voir l'influence du nombre de marqueurs et nous avons vu que même pour des marqueurs constitués d'un seul pixel, nous obtenions de bonne segmentation. Nous allons maintenant tester pour voir si la position de ces marqueurs joue aussi un rôle dans l'estimation de l'alpha matte. Pour une image de taille de 300×300 , nous avons testé le positionnement de marqueurs représentés par un unique pixel pour l'ensemble des couples possibles d'un pixel défini dans

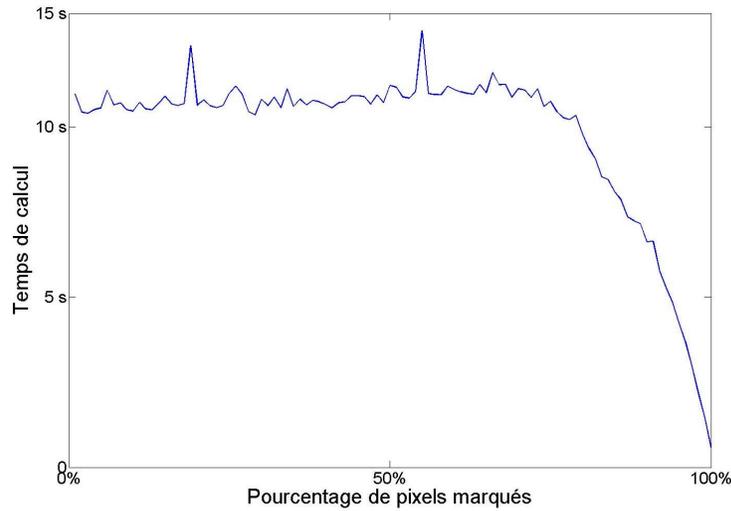


FIGURE 3.9 – Étude du "Matting" - Nombre de marqueurs - Temps de Calcul : Courbe de l'évolution du temps de calcul pour une image de résolution 300x300 pixels en fonction du pourcentage du nombre de pixels marqués.

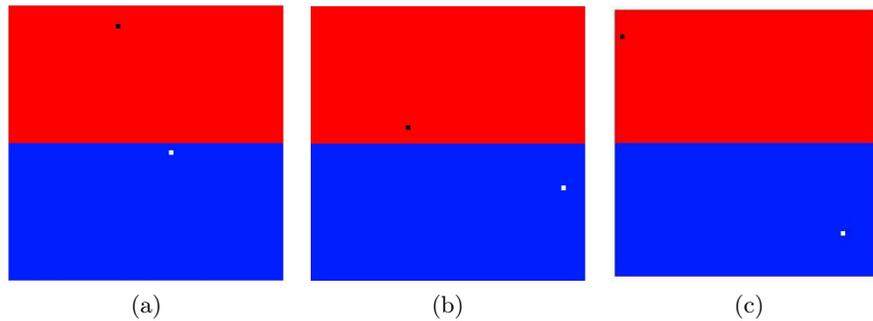


FIGURE 3.10 – Étude du "Matting" - Position des marqueurs : Exemples d'images synthétiques avec un pixels marqués positionné aléatoirement

chaque zone **F** et **B** (figure 3.10). Nous avons obtenu, à chaque fois, une bonne estimation de l'alpha matte. Ces résultats montrent que le positionnement des marqueurs n'influence pas le résultat de la segmentation. Si les marqueurs sont correctement positionnés dans la bonne zone, ils permettront d'obtenir une bonne segmentation. Ainsi le positionnement des marqueurs doit être précis dans la localisation des zones à marquer mais pas dans le positionnement à l'intérieur de ces zones. Le positionnement des marqueurs n'a pas non plus d'incidence sur le temps de calcul avec un temps moyen de calcul d'environ 14.4 secondes.

– Résolution de l'image

Nous avons vu que l'estimation de l'alpha matte nécessitait un certain temps de calcul. Ce temps de calcul peut être en partie contrôlé par le nombre de pixels marqués dans l'image. Nous avons donc aussi testé l'impact de la résolution de l'image dans l'évolution du temps de calcul. Nous avons fait varier la résolution d'image à partir d'une image de 30x30 pixels jusqu'à une taille d'environ 500x500 pixels. Pour chaque image, les marqueurs de chaque

type représentaient 50% des zones **F** et **B**. Nous obtenons un temps de calcul croissant avec l'augmentation de la résolution des images testées. Nous obtenons un temps de calcul de seulement environ 0.2 secondes pour une image de résolution 30×30 pixels jusqu'à un temps de calcul de 40.8 secondes pour une image de de 480×480 pixels. Ces résultats ont surtout mis en évidence une limite de mémoire de calcul de l'algorithme. En effet pour des images de résolutions plus grandes que 500×500 pixels, le traitement "Matting" prend trop de ressource mémoire de l'application Matlab et l'estimation de l'alpha matte échoue. Nous devons donc limiter notre étude pour des images n'excédant pas cette taille.

3.3.2 Images Naturelles

Maintenant que nous avons défini certaines limites de l'algorithme de "Matting" par marqueurs pour des images synthétiques, nous avons aussi testé l'algorithme pour des images naturelles. Nous allons tester dans un premier temps, l'influence du nombre et du placement des marqueurs pour la segmentation d'une chevelure.

Nous avons donc testé une image naturelle représentant une personne vu de face devant un fond uniforme marquée manuellement. Nous avons voulu tester ainsi l'algorithme pour la segmentation de chevelure et aussi tester l'adaptabilité de l'algorithme pour une coiffure longue où la diffusion des marqueurs doit être plus importante. Nous marquons la chevelure avec les marqueurs blancs d'information et le reste de l'image avec les marqueurs noirs de fond. Nous avons donc marqué la chevelure avec plus ou moins de marqueurs : la première où l'ensemble de la chevelure a été identifié (*figure 3.11a*), une seconde où seulement l'ensemble de la partie de la chevelure au-dessus des yeux est marquée (*figure 3.11b*) et enfin une dernière où seulement un marqueur au sommet du crâne est défini (*figure 3.11c*).

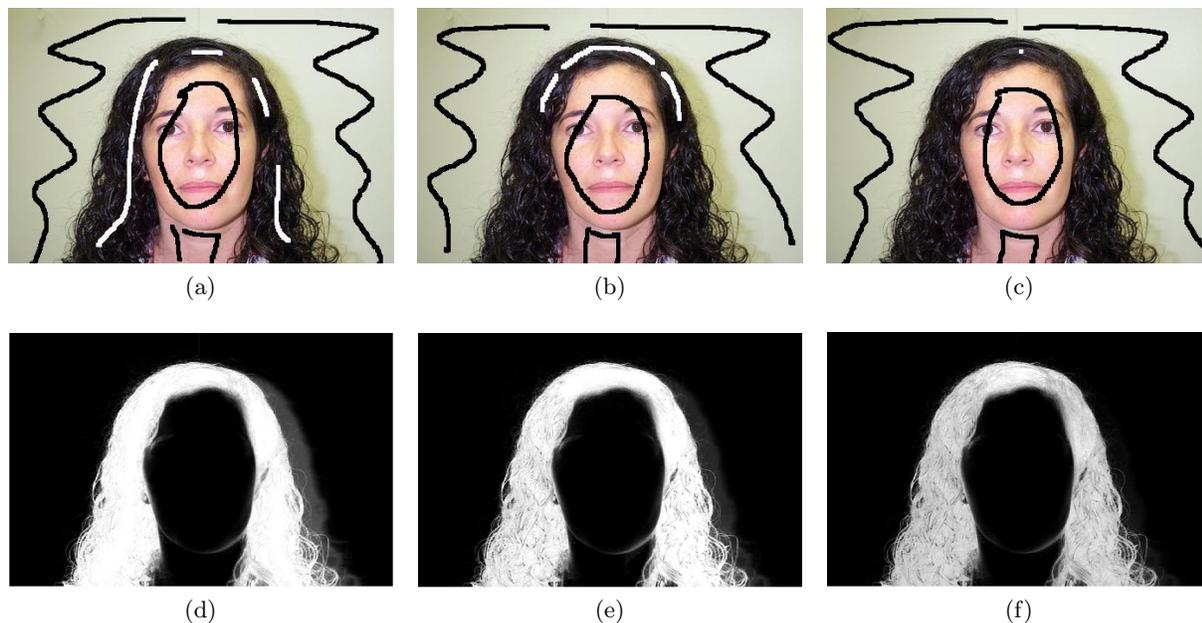


FIGURE 3.11 – Étude du "Matting" - Position et nombre de marqueurs : Exemples d'images naturelles pour tester l'influence du placement et du nombre de marqueurs cheveux nécessaire à une bonne segmentation du masque capillaire.

Nous obtenons de très bonnes segmentations pour les 3 cas d'étude. L'ensemble de la chevelure est à chaque fois correctement segmentée, seules les valeurs de l'alpha matte sont plus faibles pour les pointes de la chevelures dans le cas où seulement le sommet du crâne a été marqué (*figure 3.11f*). Ces résultats montrent la stabilité de l'algorithme de "Matting" pour les images naturelles. Ces résultats sont toutefois à relativiser, en effet même si nous obtenons de bonnes segmentations, nous admettons que l'image présente des conditions de traitement que nous pouvons considérer comme plutôt favorables : le modèle possède une coiffure assez simple d'une couleur homogène devant un fond uniforme.

Nous avons vu que pour des images naturelles, un faible nombre de marqueurs suffisait pour obtenir de bons résultats de segmentation. Nous avons voulu tester les résultats que nous pouvions obtenir pour une mauvaise définition du placement des marqueurs voire une mauvaise définition des marqueurs en oubliant de marquer certains pixels.

Dans le premier cas, la chevelure a été marquée comme zone d'information avec les marqueurs blancs et le fond avec les marqueurs noirs, seul le visage n'a pas été marqué (*figure 3.12a*). Dans le second cas, la chevelure n'a pas été marquée, seul le visage a été marqué avec les marqueurs blancs et le fond toujours avec les marqueurs noirs (*figure 3.12b*). Enfin dans un troisième cas où la chevelure a été marquée comme zone d'information avec les marqueurs blancs et le visage avec les marqueurs noirs, seul le fond n'a pas été marqué (*figure 3.12c*).

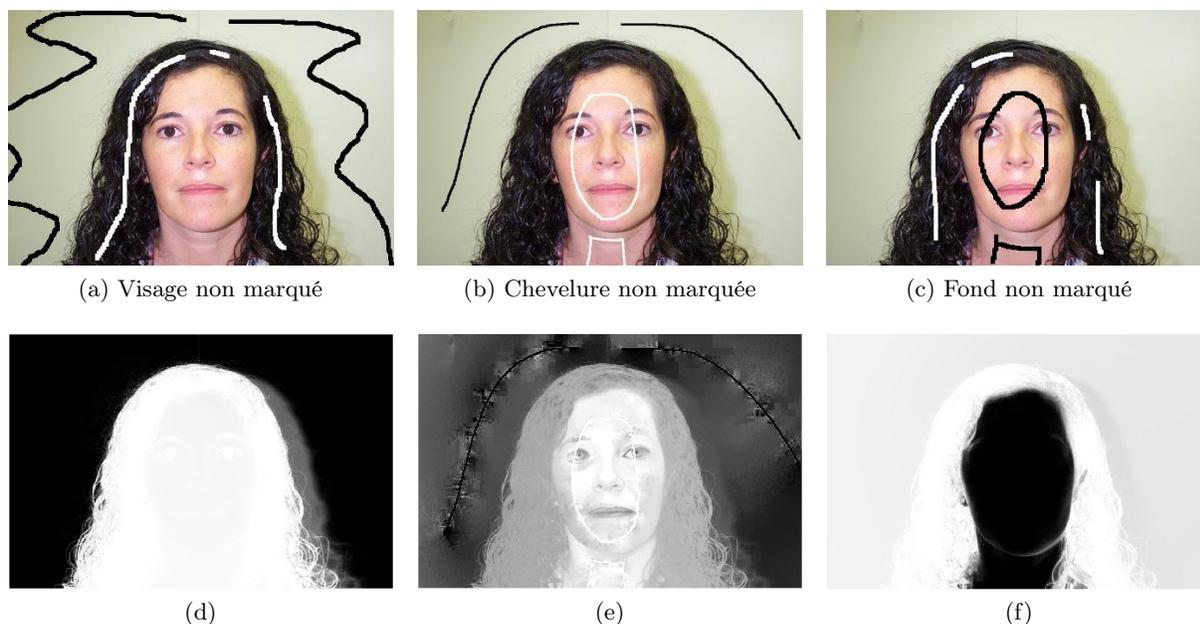


FIGURE 3.12 – Étude du "Matting" - Segmentation de zones non marquées : Les zones du visage (a) et le fond (c) sont segmentées comme de la chevelure car l'influence des marqueurs blancs identifiant la chevelure a limité celle des marqueurs noirs identifiant respectivement le fond et le visage. La segmentation de la zone de la chevelure (b) est imprécise à cause de l'influence conjointe des marqueurs blancs et noirs pour la chevelure.

Nous obtenons des résultats intéressants qui vont nous permettre de mieux comprendre les erreurs de segmentation et de déterminer les erreurs de marquage d'une image en fonction des artefacts de segmentation obtenus. Nous observons que la diffusion des marqueurs dans les zones

non marquées dépend de l'influence des autres marqueurs dans les zones voisines. Ainsi la zone du visage (*figure 3.12d*) et le fond (*figure 3.12f*) sont segmentées comme de la chevelure car l'influence des marqueurs blancs identifiants la chevelure a limité celle des marqueurs noirs identifiant respectivement le fond et le visage. Dans le deuxième cas (*figure 3.12e*), la segmentation de la zone de la chevelure est imprécise avec des valeurs de α aux alentours de la valeur moyenne de 0.5. Ces résultats s'expliquent par l'influence conjointe des marqueurs blancs et noirs pour la chevelure. Nous pouvons alors assimiler les zones où l'alpha matte prend des valeurs proches de 0.5 comme des zones d'incertitudes de la segmentation.

Maintenant que nous avons testé certaines des conditions aux limites de l'utilisation du traitement "*Matting*", il nous faut définir notre cadre d'étude pour adapter au mieux son utilisation dans le cadre de la segmentation capillaire.

3.4 Cadre d'étude

Cette thèse s'inscrit dans la continuité des recherches effectuées au sein de l'équipe AG π G du département Image & Signal du laboratoire GIPSA-lab. Depuis plusieurs années, l'équipe travaille sur l'amélioration des techniques d'analyses et de segmentations des indices visuels dans un visage [Stillitano 09] [Bouvier 10] [Rousset 08]. Nous avons aussi travaillé en partenariat avec l'entreprise Vesalis 1.1.2 [Vesalis] pour des applications de maquillage virtuel. Ces 2 axes de recherches permettent de définir un cadre et des hypothèses de recherche que nous avons suivi durant les 3 ans de recherche.

De par sa grande variabilité, tant en couleur qu'en forme, la segmentation capillaire est un problème complexe. Il a fallu développer une approche adaptative s'appuyant sur des paramètres de la chevelure présents dans la majorité des images traitées. Il a aussi fallu dans notre approche prendre en compte la possibilité de traiter une grande variété d'images prises sous de nombreuses conditions d'acquisitions différentes.

3.4.1 Formalisme

Dans cette section, nous définirons le formalisme utilisé pour appliquer le traitement "*Matting*" à la segmentation capillaire et présenterons le vocabulaire utilisé au cours de ce manuscrit.

Le but de notre algorithme de segmentation capillaire est de déterminer la position des pixels image qui correspondent à la chevelure. Nous représenterons le masque capillaire segmenté à l'aide d'une carte binaire qui se base sur la définition de l'alpha matte résultant de l'estimation du paramètre α . Pour chaque pixel, l'alpha matte définit son opacité et son appartenance au plan **F** ou **B**.

- La zone d'information ou de premier plan (*Foreground*) **F** correspond donc à la zone capillaire.
- La zone de fond (*Background*) **B** correspond à la zone non cheveux.

Bien que l'alpha matte soit défini linéairement entre $[0, 1]$, le résultat de l'estimation de l'alpha matte sera toujours considéré comme seuillé pour une valeur de $\alpha = 0.5$. Le choix du seuil est fixé arbitrairement de façon à ce que la différence entre deux segmentations du même masque capillaire ne résulte que de l'initialisation de l'algorithme de "*Matting*" par des marqueurs placés différemment.

Pour obtenir l'estimation de l'alpha matte, il a fallu initialiser l'algorithme de "Matting". Ainsi comme pour une *Trimap* on peut découper notre image en 3 zones :

- la zone Cheveux : H .
- la zone Non Cheveux : \bar{H} .
- la zone inconnue ou indéfinie : U_n .

Enfin les marqueurs H_m seront visualisés en blancs pour les pixels cheveux et les marqueurs \bar{H}_m en noir pour les pixels non cheveux

3.4.2 Hypothèse de travail

Afin de cadrer les conditions expérimentales de notre étude, nous avons défini des hypothèses de travail suivantes :

- **Modèle de coiffure** : nous avons limité notre étude capillaire seulement à l'étude de modèles féminins. Ce choix a été fait selon plusieurs critères :
 - *La diversité capillaire* : les coupes féminines sont beaucoup plus variées que les coupes masculines. De plus la majorité des coupes dites masculines se retrouvent englobés dans l'ensemble des coupes féminines.
 - *Prise en considération de la pilosité faciale* : En ne travaillant que sur des modèles féminins, nous nous affranchissons de tous les problèmes de pilosité faciale de caractéristiques très proches de la chevelure : barbe moustache ... Leurs prises en compte ultérieure sera une perspective de développement de notre approche.
 - *Prise en considération de la calvitie*. Une des difficultés majeures dans la recherche de la chevelure c'est qu'elle ne peut être absente d'un visage. Nous limiterons notre étude aux personnes non chauves afin d'être sûr de pouvoir trouver un masque capillaire dans une image. La prise en compte des personnes chauves est aussi une perspective de développement de notre approche. Cette détection pourra s'effectuer en amont du traitement afin de définir si oui ou non on doit effectuer notre algorithme de segmentation ou elle pourra s'effectuer en aval du traitement, après caractérisation du masque de segmentation qui permettra de décider si le résultat peut correspondre à un masque de chevelure.
- **Modèle de prise de vue** :
 - La pose et l'orientation du visage seront limitées à des modèles vus de face ou légèrement tournés pour éviter la perte d'informations par occultation d'une composante structurelle de visage.
 - Les occultations du visage résultant de la présence de mèche ou frange de cheveux recouvrant une partie de visage quant à elles seront prises en compte tant que les yeux ne seront pas occultés. En effet dans l'optique de l'application de maquillage virtuel, la détection des sourcils pourrait être une composante rattachée à l'analyse capillaire. Ainsi la détection des yeux est nécessaire pour dissocier les sourcils du reste de la chevelure.
 - Le contrôle des émotions du visage, des conditions de prise de vue (illuminations natu-

relles ou sous flash, images en intérieure ou extérieure) et le format de compression ne seront pas considérés comme des paramètres restrictifs.

3.4.3 Choix des bases d'images

Nous avons utilisé 3 bases d'images afin de tester différents aspects de la segmentation capillaire :

- **Base Maquillage** : Base contenant beaucoup de modèles différents de coiffure avec un fond uniforme. Cette base est utilisée pour tester notre algorithme sur la variabilité d'une chevelure
- **Base Caltech** : Base contenant peu de modèles mais chaque modèle est pris sous des conditions d'acquisitions différentes (illumination, différences de fonds). Cette base est utilisée pour tester la stabilité de notre algorithme.
- **Base Mannequins** : Base contenant de nombreuses photos de célébrités (stars du cinéma, chanteuses, mannequins) afin de réunir les paramètres des deux premières bases : une grande variabilité de chevelure et de nombreuses conditions de prises de vue.



FIGURE 3.13 – Exemple de la base Maquillage

La Base Maquillage

La base est composée de 300 images de visages de personnes toutes différentes. Les images sont de tailles uniques en 768x576 en format Jpeg. Les prises de vue représentent chaque personne vue de face devant un fond uniforme sous des conditions d'éclairages fixes. Cette base est initialement destinée pour les applications Maquillages, c'est pourquoi elle est constituée uniquement de modèles féminins présentant tout type de couleur de peau, de teinte et de texture de chevelure. Cette base comprend une grande diversité capillaire qui va nous permettre de prendre en compte de nombreuses configurations spatiales et colorimétriques de la chevelure.



FIGURE 3.14 – Exemple de la base Caltech

La base Caltech

La base est composée de 450 images de visages de 27 personnes différentes. Les images sont de tailles uniques en 896x592 en format Jpeg. Les prises de vue représentent chaque personne vue de face devant un fond varié sous des conditions d'éclairages variables. Cette base est initialement destinée pour les applications de détection et de reconnaissances faciales, c'est pourquoi elle est constituée de peu de modèle présentant peu de variation de teinte et de texture de chevelure. Cette base va nous permettre de tester la stabilité de notre algorithme pour des conditions d'acquisition peu contraintes.

La Base Mannequins

La base est composée d'environ 1000 images de visages de différentes stars du cinéma, chanteuses ou mannequins qui sont très exposées médiatiquement : leurs coiffures étant alors un ustensile de mode au même titre que le maquillage ou les vêtements. Les prises de vue sont donc très diverses autant dans le placement et l'orientation du visage dans l'image ainsi que dans la diversité des conditions d'acquisitions (fond et conditions d'illuminations). Cette base a deux objectifs :

- C'est une base hétérogène combinant les contraintes de segmentation des deux bases précédentes.
- Elle permet d'avoir différents types de coiffures pour la même personne. Son utilité était à terme de pouvoir avoir des modèles avec des coupes différentes afin de caractériser la chevelure par rapport au visage (implantation disposition). En effet la plupart des bases d'images de visages pour la reconnaissance faciale ne comportent pas assez de variation capillaire entre les différentes photos d'un même modèle.



FIGURE 3.15 – Exemple de la base Mannequins

3.5 Évaluation des résultats

La plupart des travaux présentés dans le chapitre 2 évaluent leurs résultats qualitativement. Afin de pouvoir quantifier nos résultats ainsi que l'évolution des développements mis en place durant la thèse nous avons mis en place une analyse objective de nos résultats de segmentation.

3.5.1 Base de Vérités Terrains

Nous n'avons pas trouvé de base de références capillaires pour évaluer nos résultats. La création d'une telle base est de plus, compliquée et fastidieuse. En effet la délimitation manuelle des pixels d'un masque capillaire peut être très longue et sa précision sujette à l'expertise de l'opérateur.



FIGURE 3.16 – **Limitations de la segmentation manuelle d'une chevelure** : Représentation de mèches pour lesquelles la segmentation manuelle est très délicate.

Autant pour un visage, la définition d'un masque de peau apparaît assez simple à délimiter, autant pour les cheveux, la frontière extérieure est très difficile à définir. Dans la figure 3.16, la présence de mèches très fines, voire de cheveux seuls, ne permet pas toujours une représentation pixel suffisante pour les considérer dans notre masque de référence de chevelure même si ces pixels apparaissent clairement comme des pixels cheveux pour un observateur humain.

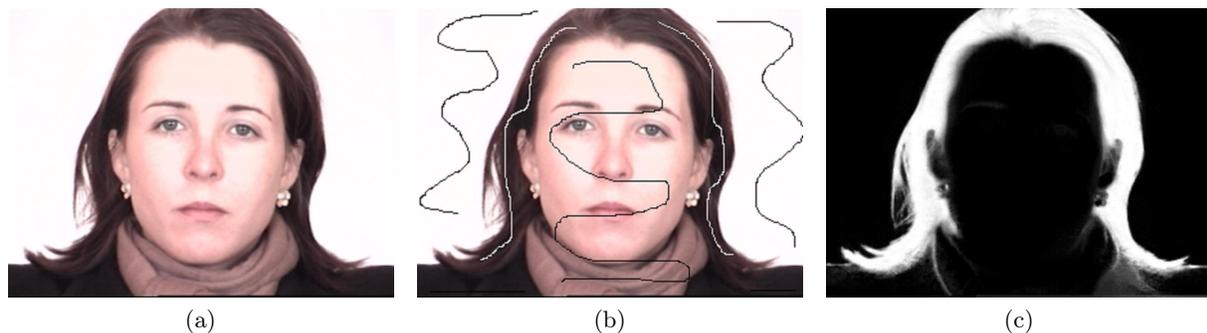


FIGURE 3.17 – **Étapes de la Création d'une Base de référence capillaire** : (a) Image Initiale, (b) Image marquée manuellement par des marqueurs (H_m en blanc & \bar{H}_m en noir), (c) Alpha matte résultat de la segmentation

Ainsi la création d'une telle base par les méthodes conventionnelles s'est vite révélée inadaptée. L'utilisation du traitement "*Matting*" pour la segmentation de la chevelure nous a permis d'envisager la création en 2 étapes d'une base de référence semi automatique en combinant l'expertise d'un opérateur et une segmentation automatique (*figure 3.17*) :

- **La première étape "manuelle"** : le positionnement des marqueurs cheveux H_m et non cheveux \bar{H}_m par un opérateur expert (*figure 3.19c*).
- **La seconde étape "automatique"** : l'estimation d'un alpha matte à partir des marqueurs afin de segmenter un masque capillaire (*figure 3.17c*).

La création d'une base par cette approche semi automatique offre plusieurs avantages :

- L'expertise de l'opérateur identifiant la position des marqueurs de la chevelure doit être moins importante. De plus la tâche est beaucoup moins fastidieuse car seulement quelques pixels (marqueurs) sont nécessaires pour initialiser l'algorithme de "*Matting*".
- Les pixels de frontière du masque de chevelure sont estimés par le traitement "*Matting*". Si dans l'ensemble, les frontières de la chevelure sont bien définies, il existe des limites à une telle approche et nous pouvons observer des artefacts de segmentation :
 - *La sous-segmentation de certaines mèches de cheveux* : la représentation de ces mèches ne comportant pas assez d'information, la diffusion des marqueurs n'a pas permis la segmentation de tous les pixels.
 - *La sur segmentation d'élément du fond similaire à la chevelure* : Certains éléments du fond peuvent avoir des caractéristiques très proches de la chevelure, ces éléments sont segmentés en même temps que le masque capillaire. Le marquage pourrait être très précis afin d'empêcher la segmentation de ces artefacts. Mais il ne nous sera pas toujours possible d'empêcher la segmentation d'artefacts trop similaire à la chevelure étudiée. L'exemple le plus concret et le plus représentatif est montré dans la figure 3.18 où une personne blonde se retrouve devant une couronne de Noël où se trouve de la paille. Même s'il est facile pour un observateur humain attentif de différencier la chevelure de la paille il est en revanche plus compliqué de la différencier par une approche numérique. Ainsi conserver cet artefact dans la définition du masque capillaire de référence n'est pas forcément une erreur rédhibitoire.

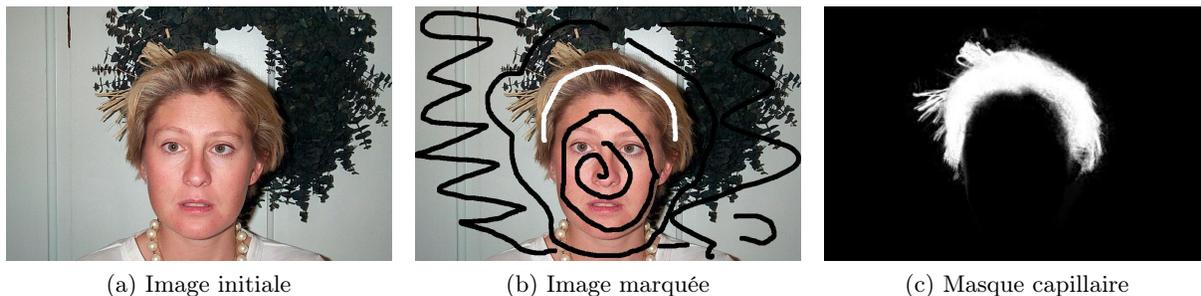


FIGURE 3.18 – **Segmentation capillaire - Cas particulier** : Exemple de sur segmentation d'artefact de la chevelure : la paille de la couronne de Noël . S'il est facile pour un observateur humain de différencier la chevelure de la paille il est en revanche plus compliqué de la différencier par une approche numérique

3.5.2 Évaluation de la segmentation capillaire

Nous évaluons nos résultats en se basant sur la détermination de deux mesures utilisées usuellement pour l'évaluation des systèmes biométriques. L'application de ce système d'évaluation pour notre algorithme est faite par analogie entre la reconnaissance d'une personne et la détection d'un pixel de cheveu. Nous allons mesurer le pourcentage de pixels cheveux non détectés par erreur et le pourcentage de pixels non cheveux détectés par erreur après segmentation des masques capillaires par le traitement "*Matting*".

- Le taux de faux rejet (T.F.R.) correspond aux pixels cheveux H segmentés comme pixels non cheveux \bar{H} . C'est-à-dire les pixels présents dans le masque de référence et qui ne sont pas dans le masque capillaire obtenu après notre traitement.
- Le taux de fausse acceptation (T.F.A.) correspond aux pixels non cheveux \bar{H} segmentés comme pixels cheveux H . C'est-à-dire les pixels qui ne sont pas dans le masque de référence et qui sont présent dans le masque capillaire obtenu après notre traitement.

Une segmentation dont le T.F.R. est faible montrera une segmentation du masque capillaire bonne et à l'inverse un T.F.R élevé montrera une sous-segmentation du masque capillaire. Pour le T.F.A., un système avec un taux élevé montrera une sur-segmentation de l'image plus le taux sera faible moins le masque sera sur segmenté.

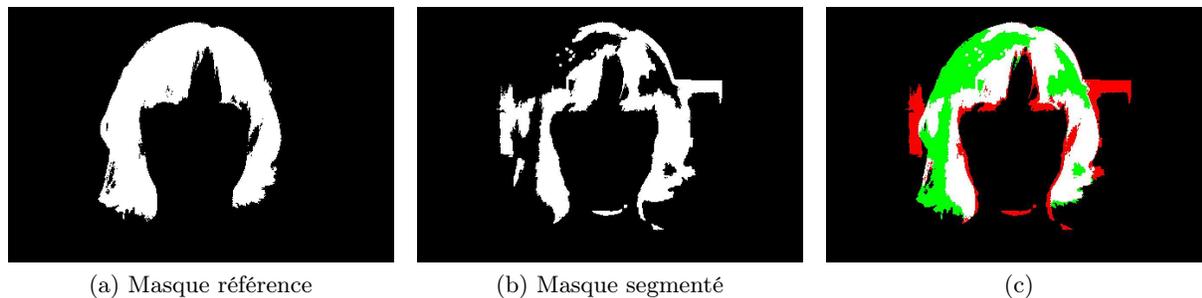


FIGURE 3.19 – **Modèle d'évaluation de la segmentation capillaire** : Représentation des pixels sur-segmentés (en rouge) correspondant au taux de fausse acceptation (T.F.A.) et des pixels sous-segmentés (en vert) correspondant au taux de faux rejet (T.F.R.)

3.6 Conclusion

Dans ce chapitre, nous avons présenté plus en détail le traitement "*Matting*". Après un rapide état de l'art, nous avons développé l'approche de Levin *et al* qui permet la segmentation d'une image à partir de l'initialisation d'un faible nombre de pixels. Nous avons testé cet algorithme sous différentes conditions pour des images synthétiques et naturelles afin de voir l'influence du nombre et du placement des marqueurs pour obtenir de bons résultats de segmentation. Nous avons aussi pu valider l'utilisation de ce traitement pour la segmentation capillaire. Cette étude, nous a permis de définir un cadre pour notre algorithme et de définir nos caractéristiques de bases d'images de test. Enfin par cette approche, nous avons pu mettre en place la création d'une base de référence qui pourra nous permettre d'évaluer nos résultats.

4 Analyse Fréquentielle



Sommaire

4.1	Analyse de texture	67
4.1.1	Approche Physiologique	67
4.1.2	Textures et chevelure	69
4.1.3	Les techniques d'analyse	70
4.2	Analyse Fréquentielle de la Chevelure	74
4.2.1	Détection de Contours	74
4.2.2	Analyse spectrale	76
4.2.3	Analyse temps/fréquence	77
4.2.3.1	Filtre de Gabor	77
4.2.3.2	Filtres par ondelettes	77
4.2.3.3	Incertitude de Heisenberg	78
4.2.3.4	Applications	78
4.2.4	Le filtre fréquentiel capillaire	79
4.2.4.1	La bande Passante σ	80
4.2.4.2	La Frequence Centrale f_0	80
4.2.5	Prétraitements	81
4.2.5.1	Apodisation	82
4.2.5.2	Correction de l'éclaircissement	83
4.3	Définition automatique des paramètres	84
4.3.1	Analyse <i>temps/fréquence</i> d'un signal une dimension	84
4.3.2	Analyse par filtres GLOP	86
4.3.3	Optimisation de la carte de détails	88
4.3.3.1	La fréquence centrale f_0	89
4.3.3.2	La bande passante σ	89
4.3.3.3	Carte fréquentielle haute et carte fréquentielle basse	90
4.4	Conclusion	92

Dans le chapitre précédent, nous avons présenté le traitement de segmentation "*Matting*" mis en place. Dans ce chapitre nous allons présenter nos travaux sur l'étude et la caractérisation de la chevelure selon son premier paramètre caractéristique : la texture. Cette première étude va nous permettre la définition d'une carte de la position probable de la chevelure afin de pouvoir positionner les marqueurs du traitement "*Matting*".

Dans un premier temps, nous ferons une étude sur l'analyse de texture : de sa définition en passant par sa représentation dans une chevelure pour finir par un court état de l'art sur les techniques existantes (section 4.1). Dans un second temps, nous présenterons plus précisément les différentes techniques utilisées pour l'analyse de texture par approche fréquentielle. Cette étude nous permettra de définir notre filtre fréquentiel de chevelure (section 4.2). Enfin dans un troisième temps, nous présenterons les différents pré traitements et traitements mis en place afin de configurer automatiquement le filtre de chevelure (section 4.3).

4.1 Analyse de texture

Avant de pouvoir caractériser une texture, il faut avant tout en avoir correctement défini la notion. Sur ce point, les définitions sont nombreuses. La définition littérale n'apporte pas vraiment d'information quant à la caractérisation d'une texture : Constitution, agencement de différentes parties d'un tout¹. Qui plus est, la notion de texture est de plus en plus utilisée pour décrire des surfaces en synthèse d'image : Image de synthèse plaquée sur un objet représenté en volume simulant un aspect de surface².

4.1.1 Approche Physiologique

Avant toute chose, afin de mieux appréhender la notion de texture pour des traitements numériques, il est intéressant de regarder l'aspect psychophysique de la perception humaine d'une texture. Une question de base qui a motivé des scientifiques à se pencher sur le sujet est : Quels sont les processus visuels mis en oeuvre afin de séparer ou segmenter deux textures ? Qu'est ce qui permet à l'homme de segmenter certaines textures et pas d'autres ? L'étude psychophysique de la perception de texture par le système visuel humain est essentielle afin de pouvoir comparer les détections faites par les algorithmes de détection ou d'analyse numériques de texture.

Le professeur Julesz, célèbre psycho-physicien qui a travaillé sur la perception du système visuel humain et en particulier la reconnaissance de forme et la perception de profondeur, est un des pionniers dans l'étude des caractéristiques des textures [Julesz 95] [Julesz 62] [Julesz 87]. D'après ses premières études, deux textures n'étaient pas différenciables immédiatement ou sans une attentive analyse, si elles avaient les mêmes statistiques d'ordre deux.

- *Les moments statistiques du premier ordre* servent à mesurer la probabilité d'observation d'informations générales sur l'intensité des pixels qui composent l'image. Les traitements se font en particulier sur l'étude de l'histogramme. La position des pixels n'intervient pas, il n'y a donc pas de notion de voisinage. La moyenne de niveaux de gris d'une image est un exemple de moment du premier ordre.

1. source l'internaute encyclopédie

2. source tout-savoir.net

- *Les moments statistiques du second ordre* servent à mesurer la probabilité de mesurer la présence d'une paire de pixels de certains niveaux de gris donnés pour une position et une orientation aléatoire. Ce sont des propriétés de paires de pixels. Les matrices de co-occurrence, par exemple, permettent de déterminer la fréquence d'apparition de "*motifs*" formés par deux pixels.

Il démontre lui-même la limite de son étude en définissant des contre-exemples comme le montre la figure 4.1.

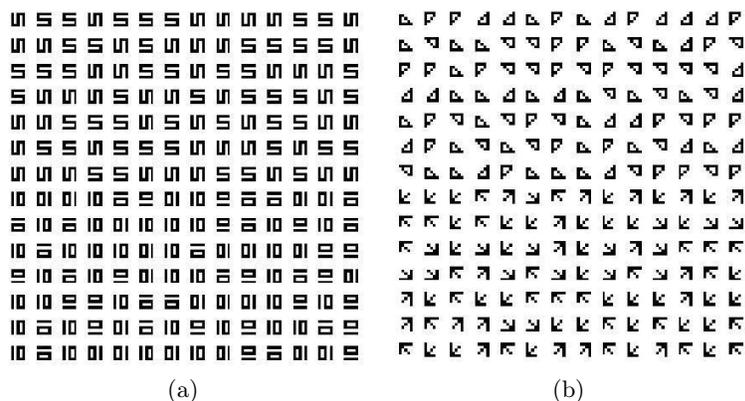


FIGURE 4.1 – "**Paire de textures**" : Deux textures qui ont les mêmes moments d'ordre deux. (a) la segmentation entre les deux régions nécessite une étude attentive. (b) la segmentation est immédiate.

L'information apportée par un pixel esseulé n'est finalement pas assez riche et discriminante. C'est pourquoi, il n'utilise plus les pixels seuls mais par l'intermédiaire d'indices visuels qui sont composés de groupes de pixels combinés au sein d'une même structure, il introduit ainsi la théorie des "*textons*". Selon cette théorie, la discrimination entre deux textures est possible si elles diffèrent en densité, en orientation et en type des simples paramètres locaux de la texture que sont les "*textons*". Les *textons* sont caractérisés par des indices visuels tels que des terminaisons, des fermetures, des points d'arrêts...

Reprenons l'exemple de la figure 4.1. Dans la figure de gauche, l'information apportée par les *textons* est la même dans la partie haute que dans la partie basse : 6 terminaisons. C'est pourquoi la détection nécessite une concentration attentive sur le motif pour pouvoir le segmenter. Par contre, pour le motif de droite, l'information apportée par les *textons* diffère, trois terminaisons pour le haut contre quatre pour le bas d'où une segmentation immédiate des deux textures.

Cette approche tend à montrer aussi la complexité de la notion de détection automatisée efficace ou non. En effet la segmentation de texture, au-delà du modèle visuel humain, ne pourra pas forcément être considérée comme efficace si elle n'établit pas une correspondance visuelle. Enfin des études en psychophysologie ont montré que le système visuel humain établissait une analyse multi-canal en fréquence et en orientation, de l'image formée sur la rétine. Les travaux de Hubel *et al* [Hubel 62], Blakemore *et al* [Blakemore 69], De Valois [Valois 82] et Harvey *et al* [Harvey 90] ont montré que le *SVH* procède à une analyse par bandes de fréquences et bandes d'orientations de l'information visuelle transmise par la rétine.

4.1.2 Textures et chevelure

La texture d'une chevelure est de manière subjective, généralement associée à la forme que prennent les cheveux. Nous pouvons donc trouver par exemple des cheveux crépus, bouclés, lisses, hirsutes ou tressés (*figure 4.2*). Nous pouvons considérer que chaque type de coiffure ne contient pas la même information de texture.

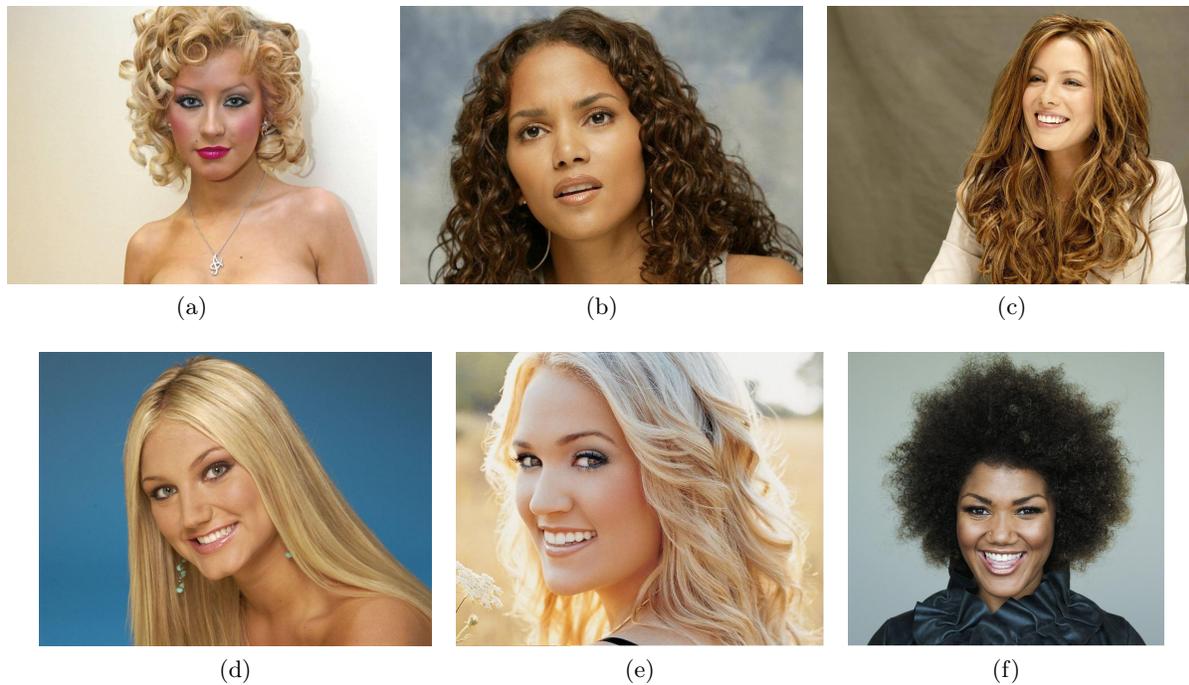


FIGURE 4.2 – **Type de texture capillaire** : Exemples de types de chevelures obtenues pour différentes formes de cheveux

La texture de la chevelure peut changer en fonction de différents paramètres :

- **Le type de cheveu** : en fonction des origines ethniques, le cheveu peut avoir une taille, un diamètre et une implantation sur le cuir chevelu qui donnent plusieurs types de cheveux qui poussent avec diverses formes. (*Annexe A.4*).
- **Le type de coiffure** : à partir de son implantation naturelle, nous pouvons changer de type de coiffure et la forme que peuvent prendre les cheveux grâce à la plasticité capillaire (*Annexe A.6.1*). Ces changements de forme peuvent être naturels ou artificiels. C'est en utilisant cette particularité du cheveu que les coiffeurs peuvent changer sa forme et créer de nouveaux types de coiffures. Au sein d'une même chevelure, on peut arriver à trouver plusieurs types de forme de cheveux. Par exemple dans la figure 4.2c nous avons des cheveux lisses sur le haut de la chevelure qui finie bouclée pour les pointes.

La texture d'une chevelure peut être changée en agissant physiquement sur les cheveux. Elle peut être aussi modifiée par sa représentation numérique dans une photo.

- **La taille du visage dans l'image** : nous travaillons sur des images numériques dont la

résolution en nombre de pixels permet d'avoir plus ou moins d'information. En fonction de la taille du modèle dans l'image à traiter, nous aurons plus ou moins de pixels décrivant une chevelure. Pour une faible résolution, l'information sera contenue dans un faible nombre de pixels où l'information sera peu discriminante. Au contraire pour des images de plus grande résolution, nous pourrions avoir de l'information pour des mèches très fines voire des cheveux uniques. Ce paramètre a d'ailleurs été une des limites pour la définition de notre base de référence (*section 3.5.1*).

- **Le volume spatial** : en fonction du type de coiffure, la chevelure peut prendre plus ou moins de volume autour du visage. En fonction de la prise de vue, l'ensemble d'une chevelure n'apparaît pas forcément dans le même plan focal. On peut obtenir une chevelure dont une partie apparaît nettement sur l'image alors qu'une autre partie apparaît totalement floue (*figure 4.2e*). Nous avons donc une perte de l'information fréquentielle puisque le flou d'une image est assimilable à un filtrage passe bas.
- **L'illumination** : de part sa structure, un cheveu est très sensible à l'illumination (*Annexe A.6.4*). Une chevelure peut ainsi avoir de nombreux reflets spéculaires où l'information de l'image est saturée ce qui entraîne donc une perte de l'information quelle contient.

4.1.3 Les techniques d'analyse

Dans la littérature du traitement d'images, on recense presque autant de définitions de textures que d'applications différentes. Cette multitude d'interprétations a même fait l'objet d'un catalogue par Coggins [Coggins 83], dont voici deux exemples :

- *"Une région dans une image a une texture constante si un nombre de paramètres locaux (structurel, statistique) de la fonction image sont constants, peu variants ou approximativement périodiques."*
- *"Une texture représente, à une échelle donnée, le même aspect quelle que soit la zone observée."*

Nous pouvons constater que la définition d'une texture dépend de la personne qui la fournit au regard de sa perception et des applications relatives à sa caractérisation. De même il existe un grand nombre de méthodes différentes pour analyser les textures qui se basent sur ces aspects. Nous allons présenter les approches les plus utilisées [Tuceryan 98] :

- **Approche Géométrique** : L'approche géométrique se base essentiellement sur les textures constituées par des 'éléments ou motifs de texture'. L'analyse s'appuie sur des propriétés géométriques appliquées à ces éléments.
- **Approche Statistique** : Un des paramètres définissant une texture est la distribution spatiale des pixels en niveaux de gris. On peut assimiler l'image à une réalisation d'un processus stochastique local et stationnaire. Pour chaque pixel de l'image, des paramètres statistiques, généralement d'ordre deux, sont estimés. On peut notamment parler des matrices de cooccurrence et des fonctions d'auto-corrélation.

Les matrices de cooccurrence sont un outil très utilisé pour l'analyse de texture. Elles se basent sur les moyennes d'espace des moments statistiques du second ordre. Elles per-

mettent de déterminer la fréquence d'apparition d'un motif formé de deux pixels séparés par une certaine distance d dans une direction particulière θ par rapport à l'horizontale.

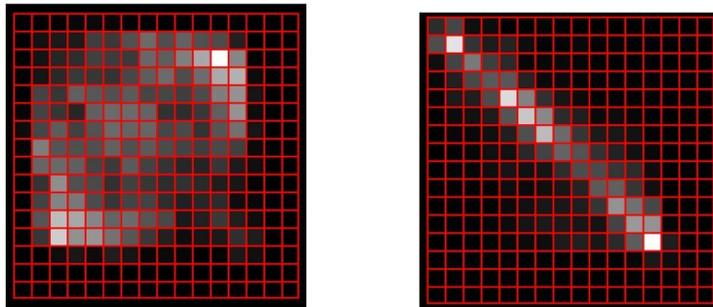


FIGURE 4.3 – **Matrice de cooccurrence** : Représentation de matrice de cooccurrence définie sur 16 niveaux de gris pour deux types de textures différentes

La matrice de cooccurrence brute contient trop d'information qui est difficilement utilisable telle quelle. Haralick a défini plusieurs indices correspondant à des caractéristiques de textures pouvant être calculées à partir de ces matrices [Haralick 79]. Ces indices bien que corrélés entre eux permettent de réduire l'information et permettent une meilleure discrimination entre les différents types de texture : on utilise principalement certains paramètres : la moyenne, la variance, l'énergie stochastique, le contraste, l'homogénéité et l'entropie.

- **Approche basée sur un modèle** : Cette méthode d'analyse de textures est basée sur la construction d'un modèle qui peut servir non seulement à analyser la texture mais aussi à la synthétiser. Les champs aléatoires de Markov (MRF), les modèles auto-regressifs (AR) et les fonctions fractales en sont les principaux exemples.



FIGURE 4.4 – **Aspect Fractal de la Chevelure** : Figure d'une chevelure à différente échelle

En observant la figure 4.4, l'aspect de la texture de la chevelure apparaît identique pour différentes échelles d'étude : une répétition de plusieurs cheveux. Cette observation ressemble à une propriété de courbes particulières : les fractales.

- **Les Fractales** :

Les fractales sont des courbes ou des surfaces qui suivent des règles déterministes ou

stochastiques impliquant une homothétie interne. Le terme « fractale » est un terme inventé par Mandelbrot [Mandelbrot 75] pour les définir :

"La notion de fractale est essentiellement due à la caractéristique que quelle que soit l'échelle à laquelle on observe la courbe, on observera toujours les mêmes détails pour autant que le nombre d'itérations soit infini ou du moins assez important pour le considérer comme tel."

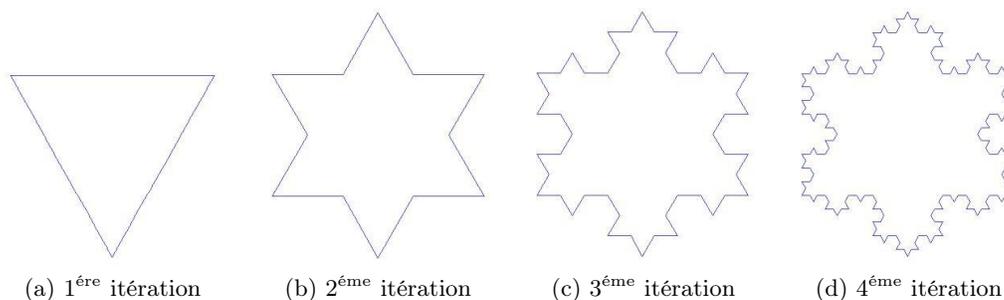


FIGURE 4.5 – **Flocon de Von Koch** : Représentation de la fractale de Von Koch. Pour chaque itération de la fractale, chaque segment est divisé en trois parties égales. La partie centrale sert de base à la construction d'un triangle équilatéral et sera finalement retirée de la figure

Un exemple classique et représentatif de la notion de fractale est la courbe de Von Koch représenté ici sous la forme dite du Flocon (*figure 4.5*). Cette courbe a longtemps été considérée comme une *"monstruosité"* mathématique. Elle s'obtient en appliquant à chaque côté d'un triangle équilatérale une transformation simple, on remplace le $\frac{1}{3}$ central par 2 segments de même longueur que le segment prélevé. Nous obtenons alors une courbe de longueur infinie : le périmètre peut être décrit par une suite arithmétique de raison $\frac{4}{3}$.

Pour caractériser ces courbes, nous utilisons la mesure de la dimension fractale, la dimension de Hausdorff D est un nombre réel positif appartenant à l'ensemble $[0, \infty[$ associé à tout espace métrique. Ainsi, contrairement aux dimensions habituelles qui sont toujours entières, la dimension fractale D peut être un nombre fractionnaire voire même un nombre irrationnel. Si un objet fractal est contenu dans un espace de dimension d , sa dimension fractale D est comprise entre 0 et d .

Pour chaque cas, la dimension D est liée au nombre d'éléments constitutants n et au rapport d'homothétie h , on obtient alors la relation définie par l'équation :

$$D = \frac{\log(n)}{\log(h)} \quad (4.1)$$

Reprenons l'exemple du flocon de Von Koch, il est formé de $n = 4$ exemplaires de lui-même réduit d'un facteur $h = 3$. Sa dimension fractale vaut :

$$D = \frac{\log(4)}{\log(3)} = 1.2618595... \quad (4.2)$$

– **Les surfaces fractales :**

Afin de caractériser la "rugosité" d'une surface, de nouvelles approches ont été développées. Si on associe une altitude à chaque niveau de gris d'une image, on obtient ainsi une sorte de relief de la surface à étudier. Nous nous sommes inspirés pour cette étude des travaux de Zhang *et al* [Zhang 05] qui ont mis en place une étude de mesure de l'usure de surfaces afin de contrôler l'usinage de disque d'aluminium. L'intérêt de cette approche est d'utiliser le spectre de la transformée de Fourier pour calculer leur mesure fractale.

Pour une image de surface en niveaux de gris $f(x, y)$ le spectre de puissance est calculé suivant l'équation :

$$S(u, v) = |F(u, v)|^2 \quad (4.3)$$

avec $F(u, v)$ la transformée de Fourier, $|\cdot|$ la norme complexe et u et v les fréquences spatiales pour les directions x et y respectivement.

Puis le spectre de puissance est converti en coordonnées polaires $Sp(f_\theta)$ avec $f_\theta = \sqrt{u^2 + v^2}$. Les valeurs de $Sp(f_\theta)$ pour chaque fréquence radiale f_θ sont ramenées à une moyenne de la distribution angulaire. Pour une surface fractale, il existe une relation linéaire entre le logarithme de $Sp(f_\theta)$ et le logarithme de la fréquence f . La dimension fractale D de la surface étudiée peut être calculée par l'équation :

$$D = \frac{(7 - \beta)}{2} \quad (4.4)$$

avec β la pente de la régression linéaire [Russ 94].

Un autre paramètre de l'étude fractale d'une surface est l'intercepte vertical de la régression linéaire [Russ 94]. L'intercepte vertical décrit la grandeur globale de la rugosité d'une surface. Il correspond à la constante de proportionnalité c dans la réécriture du spectre de puissance polaire selon l'équation [Russ 94] :

$$S(f_\theta) = c \cdot |f_\theta|^{-\beta} \quad (4.5)$$

Ce nouveau paramètre permet de différencier 2 surfaces qui peuvent avoir des propriétés différentes bien qu'ayant une grandeur fractale D identique.

Nous avons appliqué cette mesure de la dimension fractale à plusieurs images dont plusieurs échantillons de chevelure. Nous obtenons bien pour toutes ces images une dimension fractale comprise entre 2 et 3. Mais nous n'avons pas obtenu des résultats exploitables pour la segmentation capillaire. En effet pour des chevelures qui nous apparaissent similaires, nous obtenons des différences assez importantes de dimension fractale ainsi qu'à l'intérieur d'une même chevelure. Nous n'avons pas réussi à normaliser et à déterminer une mesure de comparaison pour relier l'information fractale à l'information de texture. Dans leurs travaux, Zhang *et al* [Zhang 05] étudient l'évolution de la dimension fractale d'une même pièce métallique pour mesurer l'évolution de son usure. Ils ne comparent pas au final les différentes pièces entre elles.

- **Approche par analyse spectrale :** nous avons vu que le cerveau humain décompose une image en fréquences. Ainsi les techniques de traitement du signal sont parfaitement

appropriées à l'analyse de texture par une approche fréquentielle. L'étude par analyse spectrale va nous permettre d'extraire des informations de périodicité des phénomènes qui composent le signal de l'image. On peut notamment citer les filtres dans l'espace de Fourier, les filtres multi-échelles et les filtres multi-fréquences tels que les filtres de Gabor ou les Ondelettes.

Nous avons présenté la notion de texture, abordé le cas spécifique de la chevelure et cité les principales techniques pour la caractériser. Une étude par l'approche géométrique n'est pas adaptée, en effet la chevelure ne possède pas de motifs types répartis régulièrement pour en définir des propriétés géométriques. Les approches statistiques et celles basées sur un modèle sont souvent couplées avec un découpage en blocs de l'image. Il faut donc adapter ce partitionnement de l'image à la taille de celle-ci ainsi que pour une normalisation de l'information obtenue adapter la taille des blocs à celle du visage des modèles étudiés. Ces approches sont donc assez dépendantes de la résolution de l'image. De plus après des premières études sur ces approches, il est apparu que même si nous obtenions de l'information utilisable sur les blocs, la non-stationnarité des cheveux sur l'ensemble d'une chevelure ne permettait pas une description suffisamment performante pour la définition d'une carte de localisation de la chevelure. Au final, de manière subjective, nous pouvons représenter la chevelure comme un mélange de nombreuses mèches de cheveux de plus ou moins grandes tailles répétées un certain nombre de fois. Ainsi une étude fréquentielle de la répartition de ces mèches semble la plus adaptée. De plus nous nous appuyons sur les travaux de Yacoob *et al* [Yacoob 06] qui ont caractérisé la texture de la chevelure par une décomposition par des ondelettes de Gabor. Ils obtenaient de bons résultats de classification en utilisant seulement leur descripteur de texture.

4.2 Analyse Fréquentielle de la Chevelure

L'étude par le traitement du signal va mesurer les variations de contraste de la carte de luminance. Une forte variation de contraste correspondra à une haute fréquence tandis qu'une faible variation correspondra à une basse fréquence. En effet pour une surface uniforme, il n'y aura peu voire pas de variations alors que pour une surface texturée, il y aura des variations en fonction de la "rugosité" de la texture. On peut considérer les zones homogènes d'une image comme des zones de basses fréquences et les contours comme des zones de hautes fréquences. Ainsi nous allons pouvoir analyser les différences de textures dans une chevelure et voir si nous pouvons déterminer une bande fréquentielle correspondante.

4.2.1 Détection de Contours

L'approche logique la plus simple à mettre en place est d'analyser la répartition spatiale des pixels les uns par rapport aux autres. La détection de contours par les opérateurs dérivatifs repose sur le fait que les contours correspondent à des discontinuités de la fonction de la luminance. Deux grandes approches existent : l'approche *gradient* qui calcule les extrema locaux dans la direction du gradient et l'approche *laplacien* qui détecte les passages par zéro du laplacien.

La détection de contours se divise généralement en deux étapes :

– **Définition de la carte des dérivées :**

Les filtres les plus simples sont les opérateurs basés sur les dérivées du premier ordre. On peut citer notamment les opérateurs de Roberts, Prewitt et Sobel. Ces dérivées sont calculées par convolution de l'image avec un masque de différences. Elle consiste à balayer

l'image avec une fenêtre glissante définissant des opérateurs dérivatifs. À chaque position, l'opérateur est appliqué sur les pixels de la fenêtre. Ces filtres sont très peu adaptables et sont dépendants de la taille des objets traités. Ils sont aussi très sensibles au bruit et c'est pourquoi nous conservons beaucoup de bruit haute fréquence pour les zones homogènes. Enfin ils sont anisotropes.

Un autre type d'approche repose sur la définition de critères d'optimalité de la détection de contours, ces critères débouchant sur des filtres de lissage optimaux et de gradient. Les travaux de Canny [Canny 86] sont particulièrement connus puisqu'il a défini trois critères qui permettent de bien détecter un contour :

- *Robustesse au bruit* : faible taux d'erreur dans la description de contour.
- *Bonne localisation* : le contour doit être localisé avec précision, il s'agit de minimiser la distance entre les points détectés et le vrai contour.
- *Réponse unique* : il s'agit de minimiser le nombre de réponse pour un seul contour.

Dans cette approche les contours envisagés sont modélisés par un échelon d'amplitude de type marche et le bruit est supposé blanc (de moyenne nulle). À partir de ces critères, différents filtres ont été développés qui sont plus évolués et comportent un lissage des hautes fréquences pour supprimer le bruit. Nous pouvons citer les plus connus dont notamment le filtre de Canny à réponse impulsionnelle finie, les filtres de Deriche [Deriche 87] et Shen et Castan qui possèdent une réponse impulsionnelle infinie.

– **Obtention de la carte des contours :**

Les différents filtres présentés permettent de calculer le gradient ou le laplacien d'une image mais ne donnent pas des points de contours. La segmentation dépend du type d'approche choisi. Pour les opérateurs simples, les maxima locaux du gradient sont extraits. Dans la (figure 4.6a), nous segmentons généralement que les contours forts de l'image. Nous avons bien défini le contour de la chevelure et certains éléments du visage mais nous n'avons aucune information sur la texture de la chevelure.

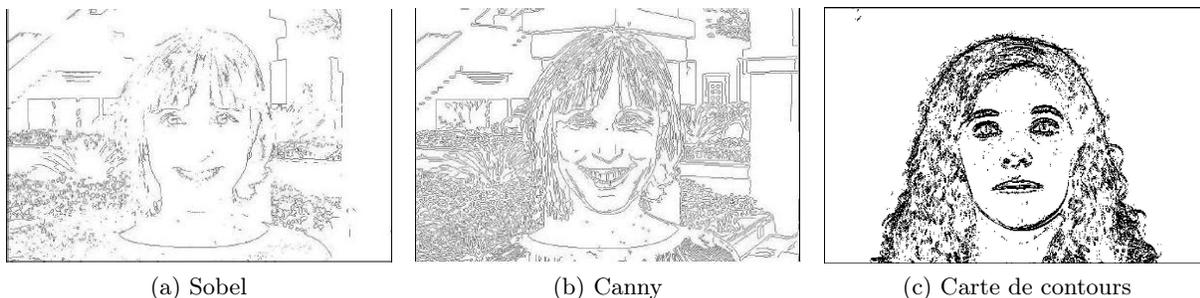


FIGURE 4.6 – **Carte de contours** : Représentation de la détection de contours par les opérateurs dérivatifs de Sobel (a), Canny (b) et par notre approche par seuillage de la valeur absolue de la carte de luminance filtrée(c)

Pour les filtres optimaux, nous obtenons une carte de contours plus précise d'autant plus que généralement elle est définie par un seuillage par hystérésis et repose sur une hypothèse de connexité. Le principe est d'utiliser deux seuils S_{haut} et S_{bas} pour la norme du gradient. Les pixels qui définissent le contour sont les pixels dont la norme est supérieure

à S_{haut} et ceux dont la norme est comprise entre S_{haut} et S_{bas} et dont un pixel voisin appartient déjà au contour. Les cartes de contours obtenues sont trop précises. Dans la figure 4.6b, nous obtenons une information de texture plus importante pour la chevelure, mais nous définissons beaucoup de contours inutiles en particulier à l'intérieur du visage.

Le but de ces filtres est de détecter et d'optimiser les contours. Au vu de la texture de la chevelure, nous ne cherchons pas seulement à détecter les contours mais aussi sa texture. En effet la chevelure est composée de plusieurs mèches. Le but n'étant pas de trouver celles qui portent le plus d'information mais d'obtenir une carte comportant le maximum de pixels cheveux.

À partir des cartes de gradient définies précédemment pour les différents opérateurs dérivatifs, nous avons redéfini une approche simple pour déterminer un seuil sur la norme du gradient. Nous calculons l'écart type de l'image filtrée et nous considérons l'ensemble des variations dont la norme est supérieure comme des zones de textures. Les pixels dont la norme est inférieure à l'écart type sont considérés comme du bruit et ne sont pas segmentés. Nous obtenons des cartes contenant plus de détails pour la zone de la chevelure tout en limitant le bruit pour le fond et les éléments du visage qui correspondent à des zones homogènes (*figure 4.6c*).

4.2.2 Analyse spectrale

Analyser une image uniquement à partir d'un opérateur différentiel spatial ne nous permet pas d'accéder à toutes les informations qu'elle contient. Afin de faire apparaître plus précisément certaines caractéristiques, il est souvent nécessaire d'en donner une autre représentation. La transformée de Fourier est un des outils majeurs dans l'analyse d'un signal. Elle permet de le décrire en fréquence ce qui apporte une information supplémentaire à la simple description spatiale.



FIGURE 4.7 – **Transformée de Fourier** : Représentation du spectre de fréquence par la transformée de Fourier d'une image.

L'analyse de Fourier permet de décrire un signal quelconque comme une somme de sinusoides de fréquences et d'amplitudes variables. Un signal est donc entièrement caractérisé par l'ensemble des amplitudes et des phases des sinusoides qui forment sa transformée de Fourier. Mais l'une des limites les plus importantes de l'analyse de Fourier est l'incapacité à localiser les zones de l'image où les fréquences sont différentes car on perd la cohérence spatiale.

4.2.3 Analyse temps/fréquence

L'analogie avec la notation musicale est communément utilisée pour introduire l'idée de représentation *temps/fréquence*, chaque note d'un morceau étant en effet associée à une certaine localisation aussi bien temporelle (son instant d'occurrence et sa durée) que fréquentielle (sa hauteur) (figure 4.8).



FIGURE 4.8 – **Représentation *temps/fréquence*** : Schéma d'une portée musicale qui modélise simplement la notion de *temps/fréquence*

On peut alors définir un outil mathématique que l'on peut comparer à une note de musique. Ainsi comme une note de musique, qui est un morceau de son apparaissant à un instant donné, d'une durée donnée et d'une hauteur donnée, nous définissons une ondelette à laquelle on peut associer des caractéristiques physiques telles que sa localisation dans le temps, sa durée et son amplitude. On peut comparer les ondelettes aux sinusoides sur lesquelles repose l'analyse de Fourier en ajoutant en plus une localisation temporelle (spatiale dans le cas d'une image).

4.2.3.1 Filtre de Gabor

Les premiers travaux ont été proposés dès le milieu des années 1940 par le physicien Gabor. Il a travaillé pour rendre l'analyse de Fourier locale à l'aide de fenêtres glissantes. Une fenêtre représente une fonction régulière et bien localisée qui est nulle en dehors de son support. En multipliant le signal étudié par une fenêtre, nous obtenons une description locale de l'information fréquentielle contenue dans le spectre de l'analyse de Fourier. Nous pouvons ainsi décrire l'ensemble de la transformée de Fourier par une collection de fenêtres de Gabor définies pour différentes fréquences et amplitudes.

Une famille d'ondelette de Gabor est définie par translation et modulation d'une fonction mère $\Phi(x)$ à support fini :

$$\Phi_{a,b}(x) = \cos(2\pi ax)\phi(x - b) \quad (4.6)$$

avec a le paramètre de modulation et b le paramètre de translation.

4.2.3.2 Filtres par ondelettes

L'analyse par ondelettes a été proposée initialement par Morlet. Elle est similaire à la décomposition de Gabor, elle décrit l'ensemble de la transformée de Fourier sous la forme d'une superposition d'une famille d'ondelettes décalées et dilatées. Elle diffère des filtres de Gabor en remplaçant le changement de fréquence par une variation d'échelle. Les ondelettes sont des copies presque identiques les unes des autres, elles sont de forme constante et ne diffèrent que par leur taille.

Une famille d'ondelette est définie par translation et dilatation d'une fonction mère $\Phi(x)$ à support fini :

$$\Phi_{a,b}(x) = \frac{1}{\sqrt{a}}\phi\left(\frac{x - b}{a}\right) \quad (4.7)$$

avec a le paramètre de dilatation et b le paramètre de translation. L'ondelette oscillera à la fréquence a pendant un temps d'autant plus court que la fréquence est haute.

4.2.3.3 Incertitude de Heisenberg

Le principe d'incertitude de Heisenberg vient de la mécanique quantique, mais il joue un rôle important dans la théorie du traitement du signal car il donne une limite à la précision de la localisation dans les différentes représentations d'un signal en se basant sur la dualité entre sa description temporelle et fréquentielle. La précision gagnée sur la description temporelle d'un signal se fait au détriment de sa précision fréquentielle. On représente l'incertitude de Heisenberg selon l'équation :

$$\sigma_t^2 \cdot \sigma_\omega^2 \geq \frac{1}{4} \quad (4.8)$$

avec σ la variance du signal respectivement dans le domaine temporelle σ_t et dans le domaine fréquentiel σ_ω .

On peut illustrer ce phénomène par l'exemple suivant : déterminer la fréquence exacte d'un signal sonore à un instant donné est impossible car pour calculer précisément sa fréquence, le signal devra être échantillonné pendant un certain temps ce qui entraînera la perte d'information sur la précision temporelle. La précision temporelle et fréquentielle d'une application seront définies en fonction du paramètre à privilégier.

4.2.3.4 Applications

Ainsi les ondelettes intègrent simultanément des paramètres de localisation et d'échelles qui sont adaptés pour de nombreuses applications en traitement d'images dont notamment la compression d'images dans une base d'ondelettes ou la détection de contours où les ondelettes sont particulièrement intéressantes car elles permettent de prendre en compte à la fois la notion d'échelle et celle de singularité.



FIGURE 4.9 – **Filtre Multi-résolution** : Représentation de la carte de détails (c) de l'image obtenue par la différence entre l'image originale (a) et une version floue (b) après application d'un filtre de lissage.

Pour la détection de contours, on utilise généralement sur les travaux de Mallat [Mallat 01] et Daubechies [Daubechies 92] sur la décomposition multi-résolution d'un signal en s'appuyant

sur l'approximation de fonctions de lissage. Pour la carte de luminance d'une image, l'application d'un filtre de lissage revient à la rendre floue et à diminuer sa résolution. Si nous considérons maintenant deux versions floues de l'image, pour des résolutions différentes, nous pouvons déterminer certains détails comme les éléments toujours présents dans la carte la moins floue mais qui ont disparu dans l'image la plus floue. Nous pouvons extraire ces détails en calculant la différence entre les deux images lissées. Nous pouvons aussi caractériser les détails en calculant la différence entre les deux fonctions de lissages qui revient au calcul d'une transformée en ondelettes.

4.2.4 Le filtre fréquentiel capillaire

En nous basant sur les différents critères de la détection de contours, de l'analyse spatiale par les opérateurs dérivatifs et par l'analyse spectrale par la décomposition en ondelette, nous avons défini notre approche pour la définition de notre carte de la localisation d'une chevelure par sa texture, nous allons combiner un filtrage fréquentiel afin de définir la bande fréquentielle qui correspond le plus à celle de la chevelure étudiée et une détection de texture par la définition d'un seuillage de la carte filtrée en analogie avec les opérateurs dérivatifs.

- Nous avons défini un filtre passe-bande isotrope représentant une couronne gaussienne dans le plan fréquence de Fourier. Il est défini selon l'équation suivante :

$$G_{(f_0, \sigma)} = \frac{1}{2\pi\sigma^2} \frac{1}{f} \exp\left(-\frac{(f - f_0)^2}{2\sigma^2}\right) \quad (4.9)$$

avec σ la bande passante et f_0 la fréquence centrale.

- *isotrope* : nous avons défini un filtre isotrope pour nous affranchir de l'information d'orientation. En effet l'important de cette nouvelle approche est surtout de définir la bande fréquentielle qui contient le plus d'information pour la meilleure détection de la texture capillaire quelle que soit sa direction.
- *Passe-bande* : nous avons défini un filtre passe-bande pour ne sélectionner qu'une bande fréquentielle précise. Nous filtrons les basses fréquences pour éliminer les composantes continues du signal et nous éliminons les très hautes fréquences afin de limiter l'influence du bruit. Nous pouvons comparer notre couronne gaussienne à la différence de 2 gaussiennes centrées de résolutions différentes. Notre filtre est alors semblable à une analyse multi-résolution par transformée par ondelettes.

Nous avons ainsi défini un filtre qui nous donnera une réponse dans une bande de fréquences particulière et selon toutes les orientations. Il est adaptable et paramétrable précisément en modifiant sa bande passante σ et sa fréquence centrale f_0 .

- Une fois l'image filtrée, nous effectuons le seuillage de la valeur absolue de la carte de luminance filtrée selon l'équation suivante :

$$\text{Carte Seuillée}(x, y) \begin{cases} = 1 & \text{si } |Imf(x, y)| > std \\ = 0 & \text{sinon} \end{cases} \quad (4.10)$$

avec std l'écart type de la norme de l'image filtrée Imf .

Nous obtenons une carte présentant de très bons résultats avec une bonne caractérisation de la chevelure en fréquence. Nous obtenons une carte dont la représentation est similaire à la

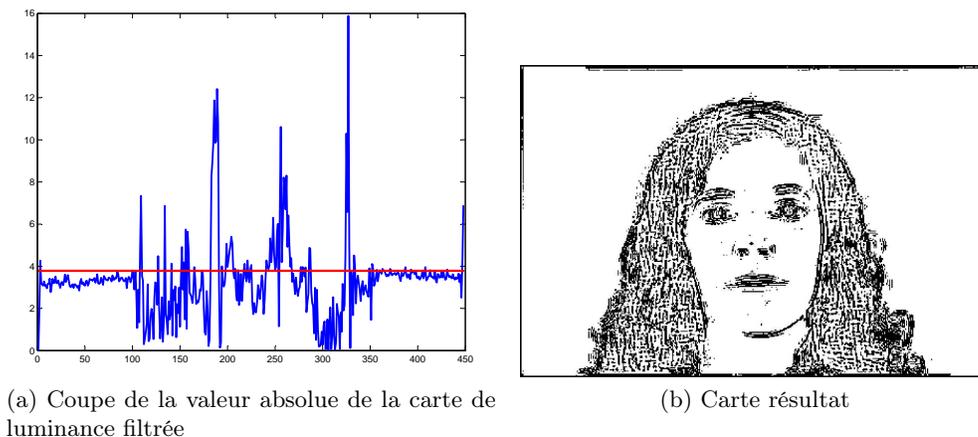


FIGURE 4.10 – **Définition de la carte de localisation fréquentielle** : Schéma de la coupe de la valeur absolue de la carte de luminance filtrée avec en rouge le seuil et la carte de localisation obtenue.

chevelure avec une représentation sous forme de mèches (*figure 4.10b*). Les résultats obtenus apparaissent meilleurs et plus fins que pour la carte de contours obtenus par l'approche gradient, avec une meilleure segmentation de l'ensemble de la texture capillaire sans augmenter la détection d'éléments à l'intérieur du visage. Pour obtenir des cartes précises, il faut effectuer un bon réglage des deux paramètres : σ de la bande passante et f_0 de la fréquence centrale.

4.2.4.1 La bande Passante σ

σ représente la largeur de la bande passante du filtre gaussien (4.9). En faisant varier σ , nous obtenons différents niveaux de filtrage.

- Pour une valeur de σ faible, nous définissons un filtre étroit fréquentiellement et large spatialement. La carte de luminance est très filtrée et nous détectons que les contours forts correspondants aux zones pour les fréquences proches de f_0 . Par exemple dans la figure 4.11a nous obtenons les contours de la fenêtre et de la porte dans le fond et une partie des contours de la chevelure.
- Pour une grande valeur de σ , nous définissons un filtre large fréquentiellement et étroit spatialement. La carte de luminance est moins filtrée, nous conservons plus d'information fréquentielle mais nous obtenons une carte de contours plus chargée car le filtre est plus sensible au bruit. Par exemple dans la figure 4.11b nous obtenons beaucoup de détails de la chevelure ainsi que du pull et des plantes du fond. On obtient aussi beaucoup d'artefacts de segmentation correspondants au bruit.

4.2.4.2 La Fréquence Centrale f_0

f_0 représente la position de la fréquence centrale du filtre gaussien (4.9). Elle prend des valeurs dans l'intervalle $[0, 0.5]$ des fréquences normalisées de la transformée de Fourier.

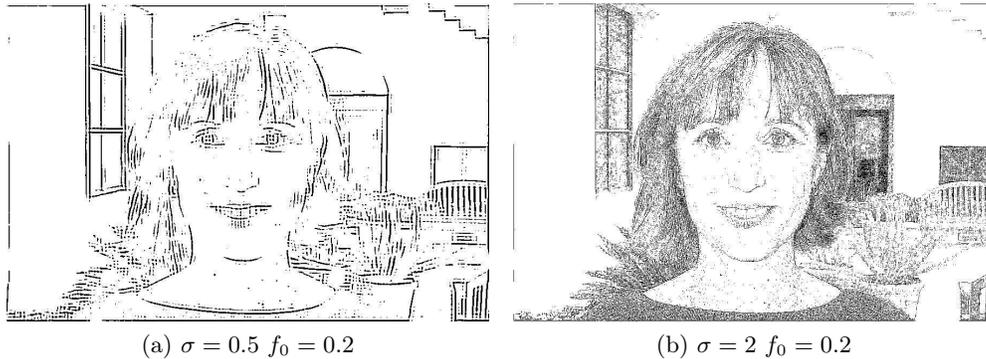


FIGURE 4.11 – **Filtre Gaussien : La bande Passante σ** : Résultats de segmentation pour différentes valeurs de la largeur de la bande de fréquence σ .

- Pour une valeur de f_0 faible ($f_0 \rightarrow 0$) c'est-à-dire pour des basses fréquences, pour une valeur de σ donnée, nous pouvons conserver une partie de la composante continue de la carte de luminance. Nous seuilons donc en grande partie l'information de luminance et seules les zones ayant une grande dynamique sont conservées. Dans la figure 4.12a, nous conservons d'une part les zones très claires telles que le mur du fond et les reflets spéculaires sur le front et d'autre part les zones très sombres telles que les vêtements, des éléments du fond comme la porte et la fenêtre et une partie de la chevelure.
- Pour une grande valeur de f_0 ($f_0 \rightarrow 0.5$) c'est à dire pour des hautes fréquences, pour une valeur de σ donnée, nous conservons certains contours mais nous obtenons surtout une carte particulièrement sensible aux bruits. Dans la figure 4.12b, nous observons que nous avons bien éliminé l'information de luminance et nous segmentons seulement de l'information bruitée très haute fréquence.

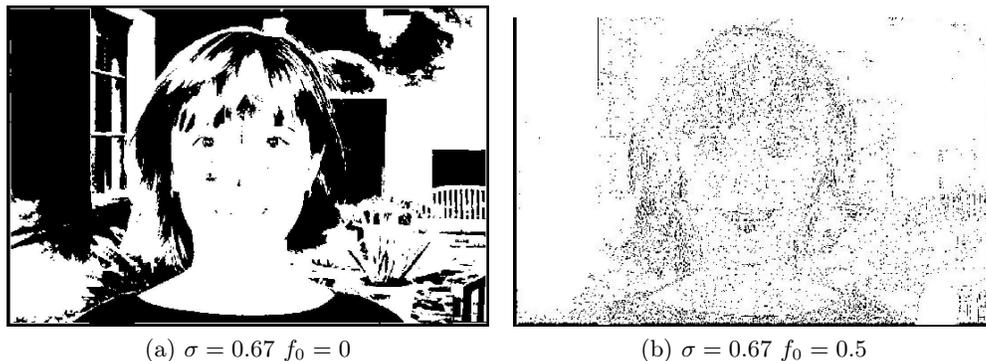


FIGURE 4.12 – **Filtre Gaussien : La Fréquence Centrale f_0** : Résultats de segmentation pour différentes valeurs de fréquence centrale f_0 .

4.2.5 Prétraitements

Afin d'optimiser la définition automatique des paramètres de notre filtre fréquentiel de la chevelure, nous présentons deux pré-traitements afin d'améliorer l'information que porte l'image traitée.

4.2.5.1 Apodisation

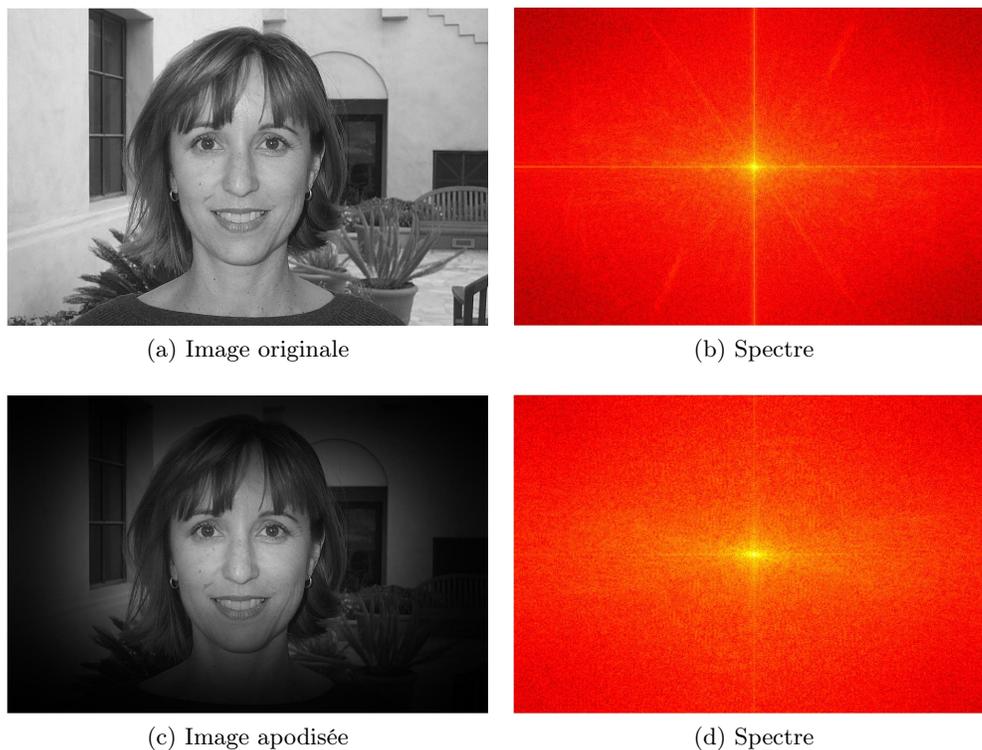


FIGURE 4.13 – **Apodisation - Résultats** : Comparaison après apodisation par une fenêtre de Hanning du spectre et de la description du spectre par la rosace de filtres GLOP.

Dans la figure 4.13b, nous pouvons observer sur le spectre une distorsion en forme de croix. Cette distorsion est due à la convolution entre la transformée de Fourier de l'image et la transformée de Fourier d'une fenêtre carré qui représente les contours bornés de l'image. Une solution pour atténuer les effets induits par la taille finie de l'image est l'application d'un fenêtrage appelé fenêtre d'apodisation. Il existe plusieurs types de fenêtres possibles : fenêtre rectangulaire, fenêtre triangulaire de Bartlett, fenêtre de Hann, fenêtre de Hamming, fenêtre de Blackman, fenêtre de Kaiser ou fenêtre gaussienne. En appliquant une fenêtre d'apodisation au spectre, la distorsion est supprimée de l'information spectrale qui est néanmoins légèrement dégradée avec un moyennage du spectre ce qui entraîne la perte légère de l'information des détails de l'image. Ce moyennage sera d'autant plus important que la taille de l'image étudiée sera petite ou que la fenêtre d'apodisation sera étroite.

Dans la figure 4.13c, nous avons appliqué une fenêtre de Hanning à la carte de luminance de l'image avant le calcul de sa transformée de Fourier. Sur le spectre obtenu, nous observons bien une atténuation des distorsions avec le retrait presque total du motif en forme de croix due au fenêtrage de l'image (*figure 4.13d*). La fenêtre de Hanning étant large, elle ne dégrade pas l'information de texture de la chevelure car en général nos modèles sont centrés dans l'image. Le moyennage du spectre est faible et ne perturbe pas la détection des détails de la texture de la chevelure.

4.2.5.2 Correction de l'éclairément

Notre étude se base sur l'étude du contraste nous allons l'améliorer pour essayer de mieux séparer les zones de textures. Nous avons mis en place un filtre de correction de la luminance basée sur la compression adaptative logarithmique des photo-récepteurs dans une rétine (*Annexe B.1*). La compression réalisée par les photo-récepteurs peut être modélisée par l'équation de Michaelis-Menten [Beaudot 94] qui à tout photo-récepteur en position spatiale p associe la luminance corrigée $r(p)$. L'équation est adaptée pour des images en 256 niveaux de gris :

$$\begin{cases} r(p) = \frac{R(p)}{R(p) + R_0(p)} \cdot (255 + R_0(p)) \\ R_0(p) = \frac{V_0}{R256} \cdot L(p) + (255 + V_0) \end{cases} \quad (4.11)$$

avec $R(p)$ la luminance d'entrée et $R_0(p)$ le coefficient qui ajuste le gain du photo-récepteur en fonction de la luminance moyenne locale $L(p)$ pour un voisinage 8-connexes. Le paramètre V_0 permet le contrôle de la compression. Ce paramètre est défini généralement pour des valeurs entre 160 (faible compression) et 250 (forte compression) [Durette 05].

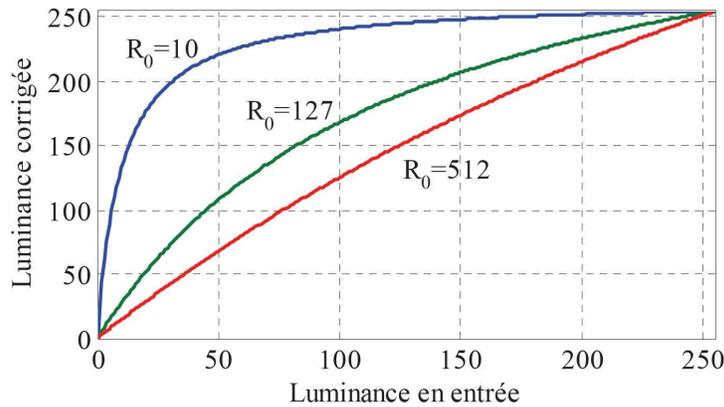


FIGURE 4.14 – **Compression adaptative logarithmique des photo-récepteurs** : Représentation de la correction de la luminance pour différentes valeurs de R_0 .

Le coefficient R_0 dépend de la luminance locale. Pour des zones sombres (avec une luminance moyenne faible), R_0 sera faible, la dynamique du photo-récepteur sera fortement augmenté. À l'inverse pour des zones claires (avec une luminance moyenne élevée), R_0 sera grand, la dynamique du photo-récepteur sera compressée. Cette compression est d'autant plus intéressante pour des images comprenant de forts contrastes et noter que la compression sera sans effet sur des images saturées.

Dans la figure 4.15, nous testons la compression logarithmique pour plusieurs valeurs de V_0 appliquée à une image où le modèle est pris en contre-jour. Pour ces zones sous exposées la luminance acquise est faible et nous perdons de l'information que le photographe percevait de ses yeux, mais n'a pu conserver sur sa photographie. En appliquant la compression logarithmique, nous corrigeons le défaut de luminance et nous restaurons l'information tant qu'elle existe (tant que les zones sombres contiennent un peu d'information sans être saturées). Dans la figure 4.15, nous avons amélioré la définition du visage qui était pris à contre-jour, nous retrouvons de l'information sur les éléments internes du visage, sur la chevelure ainsi que sur les habits. Le filtre

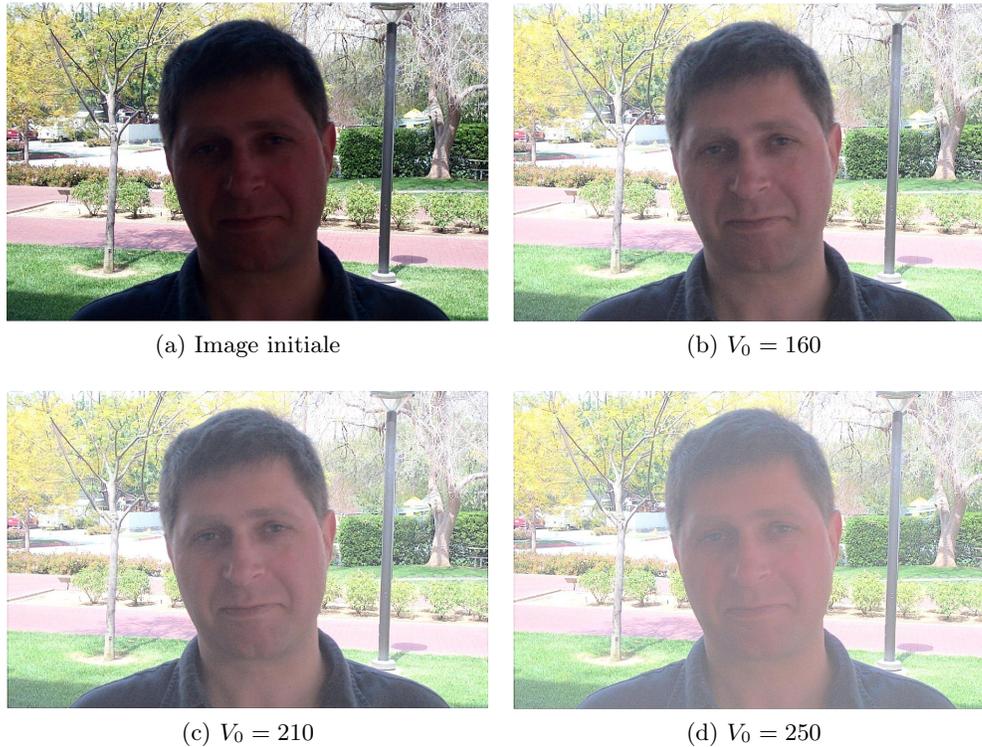


FIGURE 4.15 – **Correction de l'éclairage** : Représentation de la correction de la luminance pour différentes valeurs de V_0 .

de compression logarithmique a permis de rehausser les contrastes de l'image. Nous observons bien une meilleure séparation des couleurs dans les zones corrigées malgré le peu d'information de luminance. Nous remarquons aussi que comme prévu l'image après compression est bruitée dans les zones qui à l'origine étaient sombres. Dans la figure 4.15 en particulier pour une forte valeur de V_0 nous obtenons un bruit poivre et sel de faible amplitude sur l'ensemble des zones corrigées. Ce bruit haute fréquence pourra être éliminé par la suite.

4.3 Définition automatique des paramètres

Nous obtenons des cartes de contours très précises pour une bonne configuration de la bande passante σ et de la fréquence centrale f_0 de notre filtre. Pour notre approche de segmentation automatique du masque capillaire, il nous faut définir des règles afin de pour déterminer les valeurs de σ et de f_0 à partir d'information recueillis par l'étude de l'image traitée. Pour obtenir les meilleurs paramètres de configuration du filtre, il faut retrouver la bande fréquentielle qui correspond à celle de la chevelure.

4.3.1 Analyse temps/fréquence d'un signal une dimension

Pour notre première approche, nous sommes revenus à la base du traitement fréquentiel et nous avons testé une approche similaire à une étude d'un signal fréquentiel à une dimension. Nous ne considérons plus l'image comme un tout mais comme un ensemble de lignes (verticales ou horizontales) qui représentent chacune un signal.

La méthode la plus simple pour caractériser un signal est le calcul de sa densité spectrale de puissance (DSP). Mais cette approche n'est utilisable que si notre signal est stationnaire c'est-à-dire si ses propriétés statistiques sont indépendantes du temps. Dans notre cas, la notion de temps sera remplacée par la notion de position selon une direction dans l'image (*figure 4.16a*). Nous obtenons un signal qui décrit l'évolution de la luminosité le long d'une ligne (ou colonne) d'une image (*figure 4.16b*). Le signal étudié est non stationnaire et dépend de la position du pixel le long d'une ligne (ou colonne) de l'image. Généralement on représente ces signaux dans un repère *position/fréquence* pour montrer l'évolution fréquentielle le long d'une ligne. Nous utilisons l'analyse par spectrogramme qui est basé l'utilisation de la transformée de Fourier fenêtrée par des filtres de Gabor. Il permet de montrer l'évolution de la DSP le long d'une ligne (*figure 4.17*).

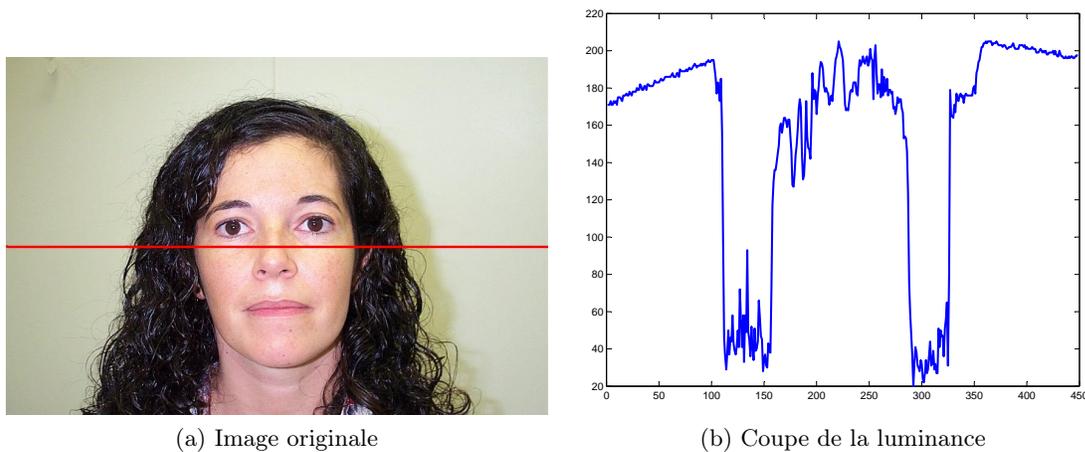


FIGURE 4.16 – **Représentation *position/fréquence*** : Schema de la coupe de luminosité décrivant la variation du contraste le long d'une ligne de l'image (b).

La figure 4.17 est le spectrogramme spatial de la ligne dont le profil est illustré dans la figure 4.16b. Nous définissons un spectrogramme spatial dont l'axe des abscisses représente l'axe des fréquences normalisées de 0 (gauche) à 0.5 (droite). L'axe des ordonnées représente l'axe des positions en parcourant la ligne de gauche (bas) à droite (haut). Nous pouvons le segmenter en 5 zones selon l'axe des positions : une zone de fond, une zone de chevelure, une zone de visage, une zone de chevelure et une zone de fond. Le fond étant uniforme il contient peu d'information fréquentielle ce qui est cohérent avec le spectrogramme obtenu où nous avons très peu d'énergie. La zone centrale correspondant à celle du visage contient aussi peu d'information avec néanmoins un peu plus d'information basse fréquence. Enfin nous observons bien les deux zones de chevelure avec une forte énergie répartie sur une large bande fréquentielle. Nous avons bien réussi à retrouver les zones fréquentielles de la chevelure et les séparer des zones plus uniformes. Cependant, comme pour l'analyse spectrale nous avons une bande de fréquence large qui ne permet pas non plus la définition d'une bande de fréquence la plus représentative de la chevelure. En général l'analyse *position/fréquence* caractérise mal les zones de fréquence de l'image car l'information de texture est trop localisée spatialement pour des signaux de quelques dizaines de pixels de longueurs seulement. Elle n'est pas assez précise pour la segmentation du masque capillaire.

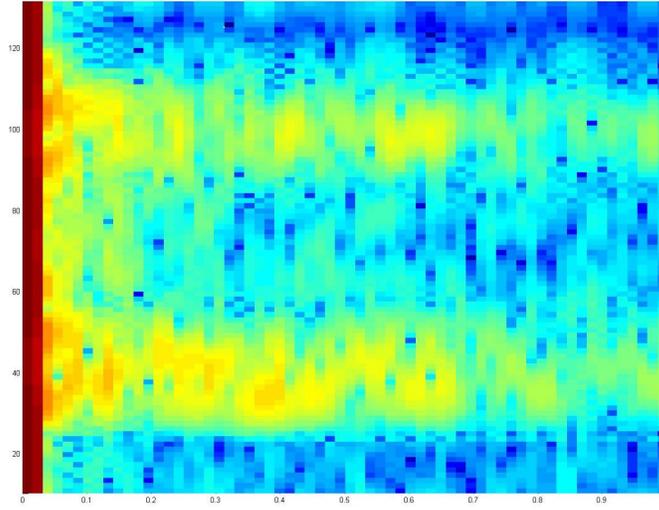


FIGURE 4.17 – **Spectrogramme** : Représentation de l'évolution de la densité spectrale de puissance le long d'une ligne. L'axe des abscisses représente l'axe des fréquences normalisées de 0 (gauche) à 1 (droite). L'axe des ordonnées représente l'axe des positions en parcourant la ligne de gauche (bas) à droite (haut).

4.3.2 Analyse par filtres GLOP

Notre seconde approche consiste à se baser sur le Système Visuel Humain (SVH) qui est un des systèmes de vision, performant. Sa modélisation est un axe de recherche important au sein du GIPSA-lab avec notamment des études sur le filtre rétinien (*Annexe B*). Le filtre rétinien procède à une analyse par bandes de fréquences et bandes d'orientations de l'information visuelle. Guyader [Guyader 06] a introduit une modélisation par des filtres de Gabor Log Polaire **GLOP** dont la modélisation est décrite par l'expression suivante :

$$G_{i,k}(f, \theta) = \frac{1}{2\pi\sigma^2} \frac{1}{f} \exp\left(-\frac{\ln\left(\frac{f}{f_k}\right)^2}{2\sigma^2}\right) \exp\left(-\frac{(\theta - \theta_i)^2}{2\sigma^2}\right) \quad (4.12)$$

Le filtre $G_{i,k}(f, \theta)$ est à variables séparables et est centré sur (f_k, θ_i) avec f_k la fréquence centrale sur l'axe des fréquences pour l'orientation θ_i . Ainsi la réponse du filtre nous donnera une énergie moyenne dans une bande de fréquences et pour une orientation particulière.

Le spectre d'amplitude de l'image est donc échantillonné par une rosace de filtres de Gabor. En prenant l'énergie à la sortie de chacun des filtres, nous obtenons une description de l'image par une matrice : *nombre d'orientation* \times *nombre de fréquence*. Cet échantillonnage réduit considérablement la dimension de l'espace de description de l'image. De plus dans cette représentation log-polaire, un zoom correspond à une translation horizontale de l'image du spectre et une rotation correspond à une translation verticale. Le simple repérage d'une translation en deux dimensions peut permettre de reconnaître la signature spectrale d'un objet zoomé et tourné.

Nous avons dans un premier temps appliqué le filtre rétinien sur l'ensemble d'une image pour observer l'information fréquentielle portée par une image naturelle. Nous avons choisi une image avec un fond chargé avec par exemple la présence de plante et de banc (*figure 4.18a*). Nous avons aussi choisi une image avec un modèle devant un fond uniforme pour prendre en compte uniquement l'information fréquentielle portée par la chevelure et le visage (*figure 4.18c*). Pour

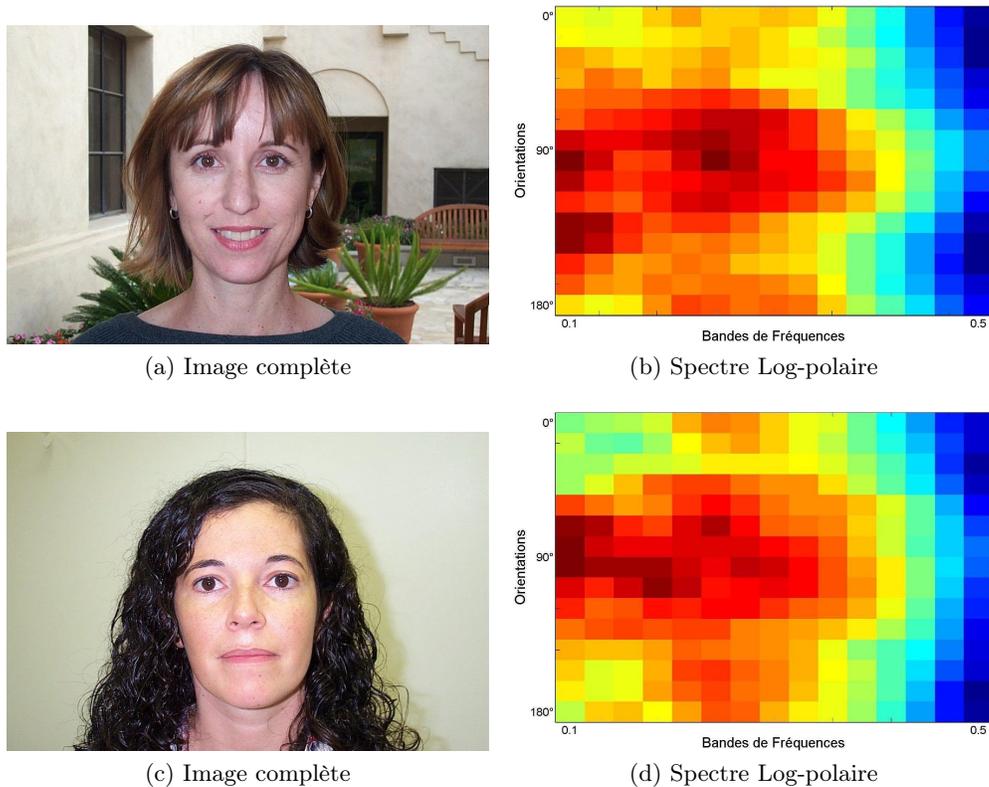


FIGURE 4.18 – **Filtres GLOP - Image complète** : Description d’images par la rosace de filtres GLOP

les 2 images, nous obtenons des résultats similaires, une répartition sur l’ensemble des fréquences et des directions (*figure 4.18*).

Nous observons que nous avons des plus fortes énergies pour les valeurs centrales de la décomposition en directions (axes des ordonnées) pour des angles autour de 90° . Nos images comportent donc plus d’informations horizontales. Ces résultats sont logiques au vu de la description des images traitées : elles contiennent entre autre des arrêtes de murs, des barreaux de bancs verticaux et des chevelures longues qui tombent le long du visage. Par contre, pour une direction donnée, nous pouvons observer un étalement du spectre pour l’ensemble des fréquences (axes des abscisses) avec moins d’énergie pour les très hautes fréquences $f \simeq 0.5$. Ces résultats correspondent à ce que nous pouvons trouver dans des images naturelles qui contiennent beaucoup d’information fréquentielle avec de nombreux éléments de différentes tailles et directions. Cependant, ce qui est à noter c’est que nous obtenons des résultats similaires pour une image avec ou sans un fond chargé. Le visage ne comportant pas a priori beaucoup d’information fréquentielle cela implique qu’une chevelure contient effectivement une large bande de fréquence.

Nous avons donc testé des échantillons de chevelure pour étudier particulièrement leur répartition fréquentielle (*figures 4.19a 4.19c*). Les résultats obtenus confirment nos premières observations, la chevelure a une bande fréquentielle large. Nous obtenons des résultats semblables aux spectres obtenus pour une image entière, nous observons bien une répartition d’énergie sur l’ensemble du spectre. En fonction du type de coiffure, nous pouvons remarquer une direction

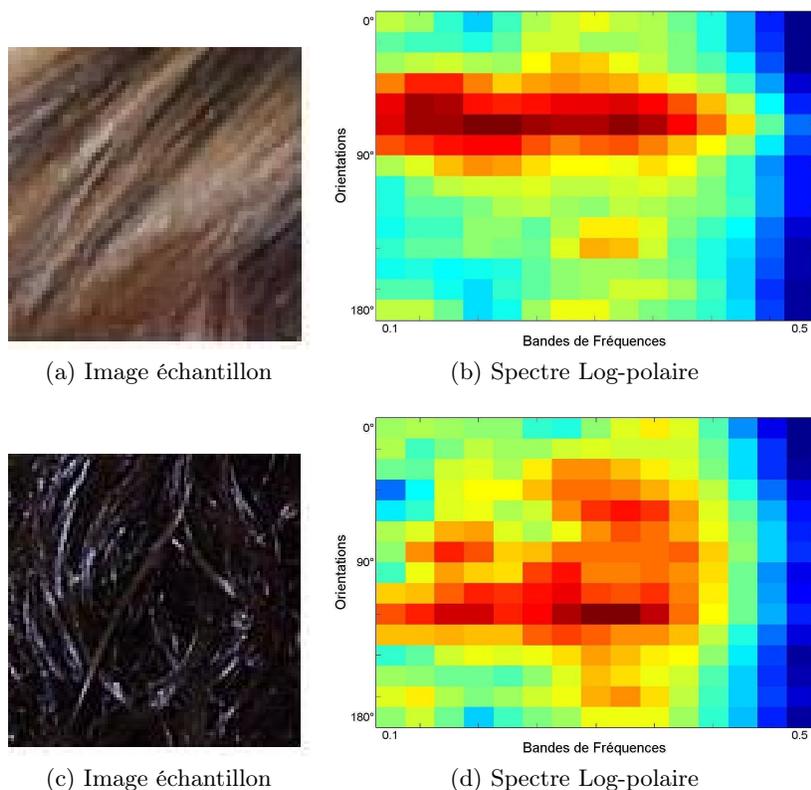


FIGURE 4.19 – **Filtres GLOP - Echantillon de cheveux** : Description du spectre par la rosace de filtres GLOP

particulière (*figure 4.19b*) ou une répartition sur l'ensemble des directions (*figure 4.19d*). Cette information de direction pourra être utilisée pour la caractérisation de la chevelure (*section 7.1*). Par contre, les différents spectres décrits par les filtres GLOP ne sont pas assez homogènes à cause de la non-stationnarité des cheveux sur l'ensemble d'une chevelure, pour caractériser différents types de coiffures et nous permettre un apprentissage par la définition de modèles de la répartition spectrale de la texture de la chevelure.

4.3.3 Optimisation de la carte de détails

Nos deux approches pour paramétrer automatiquement le filtre fréquentiel de la chevelure n'ont pas permis la définition d'une bande de fréquence particulière à la chevelure. En définissant notre filtre nous nous sommes appuyé sur l'analogie avec l'analyse par ondelette et la définition d'une carte de détails relative à une certaine résolution de l'image. Contrairement à la détection de contours par les opérateurs dérivatifs qui ne recherchent que les contours forts de l'image, nous essayons de conserver au maximum l'information de texture de la chevelure et donc une information fréquentielle plus importante. Ainsi dans cette section, nous avons cherché à optimiser la définition des zones de détails en étudiant la dynamique de la carte de luminance après filtrage.

Nous obtenons une carte de la luminance filtrée avec une alternance de variations positives et négatives centrée autour de sa valeur moyenne μ (*figure 4.20*). En fonction du filtrage défini par les deux paramètres σ et f_0 , nous filtrons plus ou moins en fréquence et en amplitude la carte

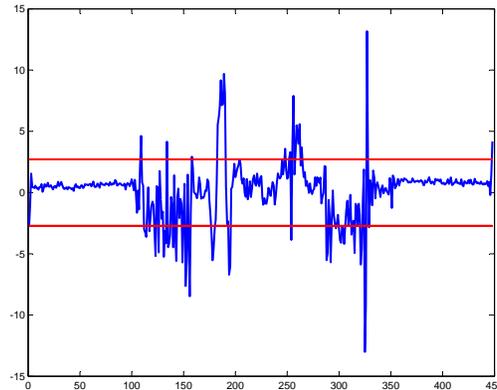


FIGURE 4.20 – **Carte de la luminance filtrée** : Coupe de la carte de la luminance filtrée centrée autour de sa valeur moyenne μ avec en rouge les 2 seuils $+std$ et $-std$

de luminance. En étudiant le rapport entre la moyenne μ et l'écart type std de l'image filtrée, nous avons défini des règles pour obtenir une bonne carte de segmentation de la position de la texture de la chevelure.

4.3.3.1 La fréquence centrale f_0

Le paramètre f_0 est déterminé automatiquement en faisant le rapport de la taille du visage par rapport à la taille de l'image. Pour une grande taille de visage nous définissons une faible valeur de f_0 et inversement, plus le visage sera petit dans l'image, plus la valeur de f_0 sera grande. Ainsi nous projetons l'intervalle $[1,5]$ du rapport de la taille du visage sur la taille de l'image sur l'intervalle $[0.1,0.5]$ des fréquences normalisées. Nous fixons la valeur de la borne inférieure de l'intervalle fréquentiel à 0.1 pour ne pas ne pas prendre en compte les très basses fréquences. Pour notre algorithme de segmentation capillaire, nous travaillons sur des images de type portrait où le modèle et sa chevelure doivent avoir une certaine taille dans l'image. En fait le rapport de taille est calculé comme le rapport moyen entre les valeurs horizontale et verticale de la taille du visage par rapport à celle de la taille de l'image. C'est pourquoi nous avons limité les valeurs du rapport de taille pour un intervalle avec une borne supérieure avec une valeur égale à 5. Pour des rapports de taille supérieure, la taille du visage dans l'image ne sera pas suffisamment grande pour avoir une bonne précision. De plus nous avons montré que l'algorithme de "Matting" nous limitait dans la taille des images que nous pouvons segmenter (*section 3.3.1*).

4.3.3.2 La bande passante σ

Une fois la valeur de f_0 fixée, nous allons chercher à adapter la valeur de σ qui permet d'obtenir la meilleure carte de segmentation fréquentielle de la chevelure. Comme nous l'avons décrit dans la section 4.2.4.1, le paramétrage de σ va nous permettre de contrôler la quantité d'information fréquentielle filtrée. Les variations du paramètre σ de notre filtre fréquentiel va nous permettre de contrôler les variations des paramètres spatiaux de la carte de luminance filtrée : sa moyenne μ et son écart type std .

- Pour un filtre peu sélectif (valeur de σ grande) : nous filtrons peu la carte de luminance et nous conservons une partie de la composante continue, nous avons donc une valeur de la moyenne μ importante. Nous conservons aussi une large bande fréquentielle, nous avons donc les alternances correspondantes à plusieurs fréquences et nous avons une valeur de l'écart type std grande.
- Pour un filtre sélectif (valeur de σ faible) : nous filtrons fortement la carte de luminance et nous éliminons la composante continue, nous avons donc une valeur de la moyenne μ qui tend vers une valeur nulle. Nous filtrons une bande fréquentielle étroite, nous avons donc les alternances correspondantes à une fréquence précise et nous avons une valeur de l'écart type std plus faible.

Le réglage de σ influe l'évolution des valeurs μ et std . L'idée est alors de faire varier σ afin de trouver la bande fréquentielle qui permette de conserver le maximum d'information de texture de la chevelure tout en minimisant le bruit. En analysant l'évolution de la valeur moyenne μ et de l'écart type std en fonction de σ , nous observons que les deux courbes décroissent pour des valeurs de σ de plus en plus faibles (*figure 4.21*) jusqu'à leur intersection en un point. Si nous calculons la carte fréquentielle de chevelure en ce point, nous obtenons une carte assez précise avec une bonne représentation de la texture capillaire. Nous pouvons alors considérer l'intersection des deux courbes comme un point remarquable et définir une heuristique de paramétrisation automatique de notre filtre fréquentiel. Nous admettons que la définition de nos paramètres n'est pas une approche optimale et que nous pouvons certainement trouver un meilleur couple (f_0, σ) qui permettent d'obtenir une meilleure carte de texture de la chevelure mais la carte obtenue est largement suffisante pour obtenir une carte de la localisation de la texture d'une chevelure utilisable pour une initialisation correcte des marqueurs pour le traitement "*Matting*". Nous avons alors mis en place une boucle itérative pour paramétrer le réglage de σ :

- Pour $\mu > std$: La valeur moyenne de la luminance filtrée est grande devant l'écart type. Pour la prochaine itération, la valeur de σ sera augmentée.
- Pour $\mu < std$: La valeur moyenne de la luminance filtrée est petite devant l'écart type. Pour la prochaine itération, la valeur de σ sera diminuée.
- Pour $\mu = std \pm \epsilon$: La boucle itérative est stoppée.

Nous avons ainsi mis en place une méthode robuste qui fonctionne pour l'ensemble des images traitées. Quelle que soit l'image filtrée, nous retrouvons toujours la condition d'arrêt de la boucle itérative. Notre approche est de plus adaptative et permet une définition de la carte fréquentielle de chevelure en fonction de l'image traitée.

4.3.3.3 Carte fréquentielle haute et carte fréquentielle basse

En observant la carte de luminance filtrée obtenue (*figure 4.20*), nous voyons qu'elle n'est pas symétrique par rapport à sa valeur moyenne. Nous pouvons trouver des zones où il y a plus d'alternances positives que négatives et inversement. Cette dissymétrie étant enlevée une fois la valeur absolue de la carte de luminance filtrée calculée pour la définition de la carte de segmentation. Néanmoins, nous avons voulu voir l'information portée par les zones positives et négatives.

- **Définition des cartes**

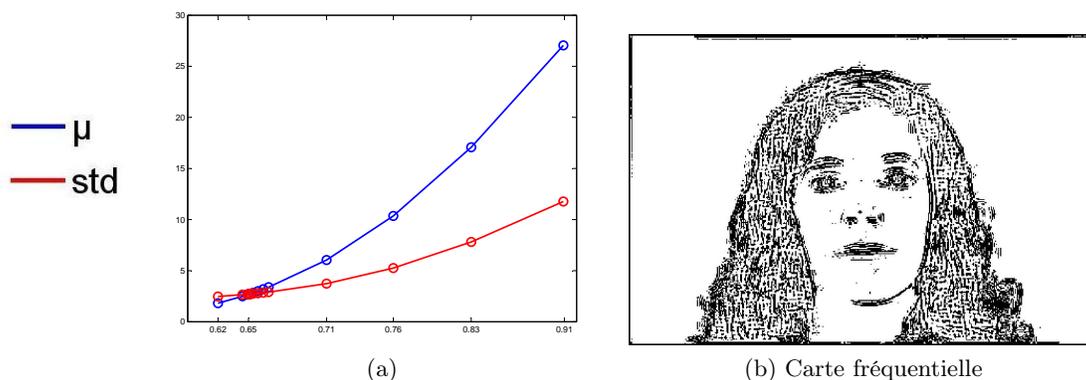


FIGURE 4.21 – **Paramétrage de la bande passante σ** : (a) Courbe de l'évolution des valeurs de la moyenne μ et de l'écart type std de la carte de luminance filtrée en fonction σ , (b) Carte fréquentielle de chevelure obtenue pour la valeur de σ définie au point d'intersection

Nous avons défini un nouveau modèle de segmentation selon les équations suivantes :

$$\begin{aligned}
 \text{Carte Haute}(x, y) & \begin{cases} = 1 & \text{si } Imf(x, y) - \mu > std \\ = 0 & \text{sinon} \end{cases} \\
 \text{Carte Basse}(x, y) & \begin{cases} = 1 & \text{si } Imf(x, y) - \mu < -std \\ = 0 & \text{sinon} \end{cases}
 \end{aligned} \tag{4.13}$$

avec std l'écart type de l'image filtrée Imf .

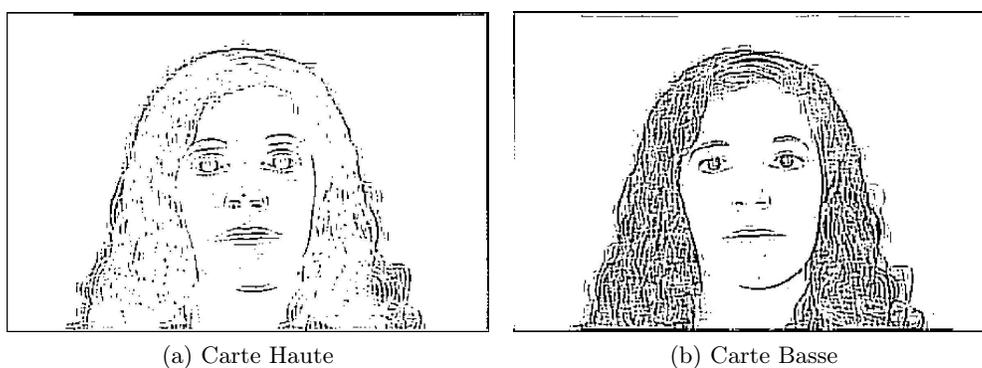


FIGURE 4.22 – **Carte fréquentielle haute et basse de la chevelure** : Représentation de la carte fréquentielle haute correspondant aux alternances positives et de la carte fréquentielle basse correspondant aux alternances négatives

Nous définissons donc deux cartes : une carte fréquentielle haute (figure 4.22a) et une carte fréquentielle basse (figure 4.22b). Nous observons qu'une seule des deux cartes est vraiment représentative du masque capillaire et que l'autre carte comporte beaucoup moins d'information et plus d'artefacts de segmentation. La carte fréquentielle définie auparavant (section 4.2.4) sur la norme de la carte de luminance filtrée est retrouvée en effectuant la somme entre les deux cartes (figure 4.21b). En général il apparaît qu'en fonction de la teinte de la chevelure, nous pouvons avoir une carte fréquentielle de la chevelure plus pré-

cise en choisissant soit la carte des alternances positives ou celles des alternances négatives.

– **Sélection de la carte**

Pour choisir celle qui représente la meilleure définition du masque de la chevelure, nous faisons une étude locale de la répartition des pixels segmentés. Nous avons connaissance de la position du masque de peau par un masque binaire obtenu après segmentation des pixels de couleur de peau (*section 5.3*). Ce masque va nous permettre de comparer la répartition des pixels fréquentiels à l'intérieur et à l'extérieur du masque de peau. Pour optimiser notre choix, nous avons limité notre analyse pour une fenêtre d'étude située au sommet du visage à partir de données géométriques de la détection du visage par l'algorithme de Viola & Jones [Viola 04] ce qui permet de ne pas prendre en compte les pixels du fond de l'image.

Nous cherchons à maximiser le nombre de pixels situés hors du masque de peau tout en minimisant le nombre de ceux situés dans le masque. En général si le paramétrage du filtre fréquentiel de chevelure a été correct, le nombre de pixels dans le masque visage est faible, la surface du front contenant peu d'information fréquentielle. Nous choisirons la carte dont le nombre de pixels hors de la carte de visage est maximum.

4.4 Conclusion

Dans ce chapitre, nous avons introduit nos travaux sur l'analyse fréquentielle de la chevelure. Après un court récapitulatif de l'analyse de texture, nous avons montré les caractéristiques de la représentation de la texture d'une chevelure dans une image, ce qui nous a permis de définir un cadre d'étude. Nous avons choisi une approche par analyse spectrale conjointement dans le domaine spatial et fréquentiel. Cette analyse nous a permis de définir notre filtre fréquentiel de chevelure modélisé par un filtre passe-bande isotrope gaussien. Ce filtre va nous permettre d'obtenir une carte des détails de texture similaire à la chevelure. Pour le paramétrer automatiquement, nous devons définir la bande fréquentielle la plus représentative correspondante à la description d'une chevelure. Nous avons essayé de la définir par une étude spectrale par l'utilisation d'un filtre rétinien et par une étude spatiale par une analyse *position/fréquence*. Ces deux études n'ont pas permis une définition précise, la chevelure ayant une large bande fréquentielle. Finalement nous définissons les paramètres σ de la bande passante et f_0 de la fréquence centrale du filtre par une optimisation de l'information sur la carte de luminance filtrée. Nous avons défini des règles basées sur la taille du visage dans l'image et combinées avec le rapport entre la moyenne μ et l'écart type *std* de la carte de luminance filtrée.

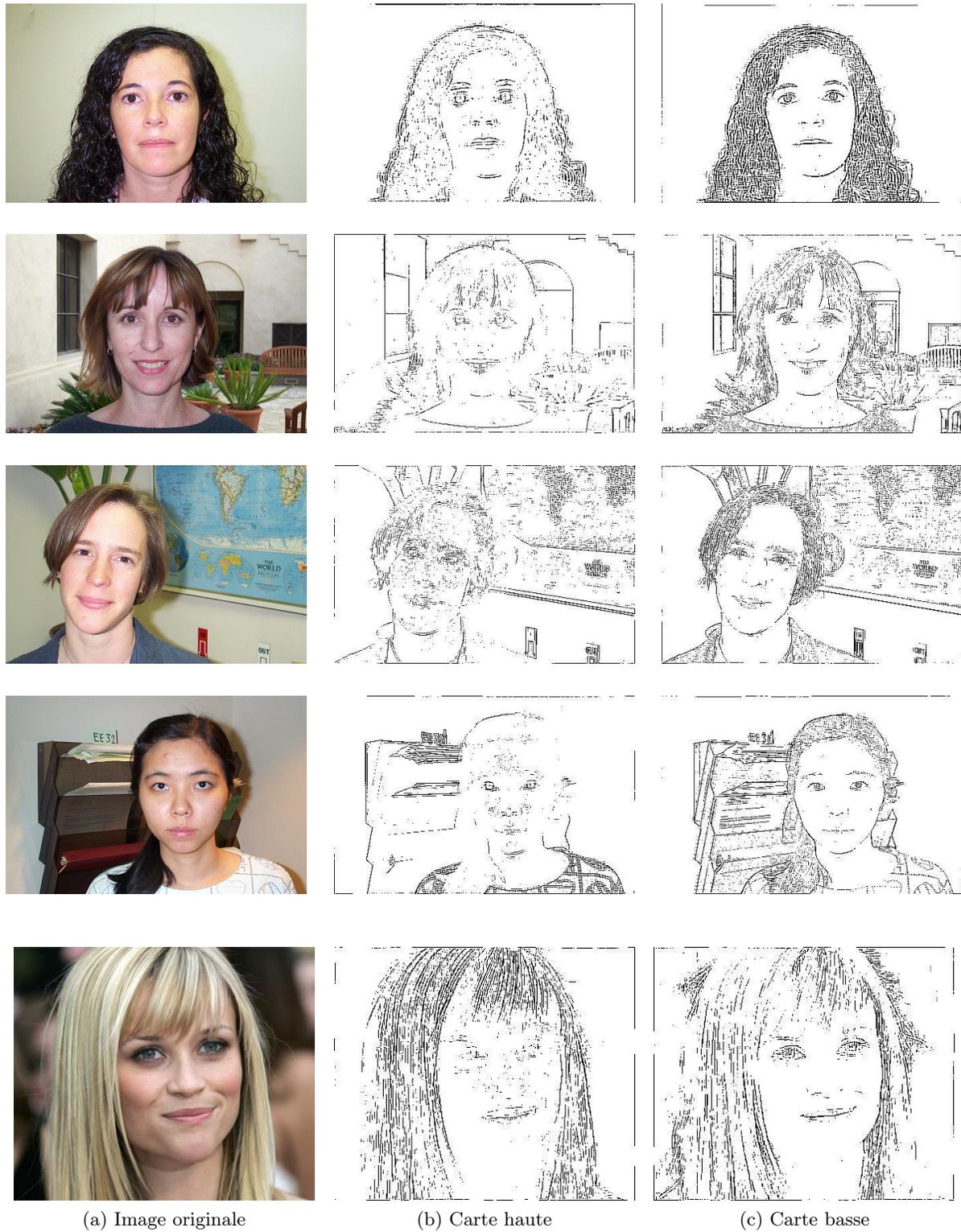


FIGURE 4.23 – **Analyse Fréquentielle - Résultats** : Exemples de cartes fréquentielles de chevelure haute et basse

5 Analyse Couleur



Sommaire

5.1	Analyse Couleur	97
5.1.1	Teintes et Chevelure	97
5.1.2	La segmentation couleur	99
5.1.3	Base de référence de masque de peau	100
5.2	La Segmentation Couleur supervisée	101
5.2.1	Le choix d'un espace couleur	101
5.2.2	La collecte d'une base d'image de référence couleur	104
5.3	Segmentation couleur de la peau	106
5.3.1	Classifieur couleur par valeurs fixes	107
5.3.2	Classifieur Couleur par Minimisation de la distorsion de couleur	109
5.3.3	Classifieur Couleur Probabiliste	111
5.3.4	Classifieur Couleur de Peau	114
5.4	Segmentation couleur de la chevelure	115
5.5	Conclusion	116

Dans le chapitre précédent, nous avons étudié le premier paramètre représentatif de la chevelure, sa texture par une analyse fréquentielle. Dans ce chapitre nous présenterons nos travaux sur l'étude et la caractérisation de la chevelure selon le deuxième paramètre caractéristique qu'est sa teinte. Cette seconde étude va nous permettre également la définition d'une carte de la position probable de la chevelure afin de pouvoir positionner les marqueurs du traitement "*Matting*". Dans un premier temps, nous ferons une étude sur l'analyse colorimétrique. Nous verrons sa représentation dans une chevelure et nous présenterons les deux grands types de segmentation couleur par approche supervisée et non supervisée (*section 5.1*). Dans un second temps, nous présenterons en détails la segmentation couleur par approche supervisée par la modélisation d'un modèle couleur (*section 5.2*). Puis nous présenterons différentes approches de segmentation couleurs en s'appuyant sur une étude bibliographique des techniques de segmentation de la peau (*section 5.1.2*). Enfin nous appliquerons la segmentation couleur à la chevelure pour la définition du masque de localisation capillaire par sa teinte (*section 5.4*).

5.1 Analyse Couleur

5.1.1 Teintes et Chevelure

Dans leur travaux, Zang *et al* [Zhang 08] et Lee *et al* [Lee 08] ont défini 4 (respectivement 5) classes couleurs afin de créer des couleurs de références pour leurs modèles d'apprentissage de leurs algorithmes de segmentation couleur de la chevelure.

Mais une chevelure peut prendre de nombreuses et diverses teintes : Aile de corbeau, Auburn, Blanc, Blond, Blond vénitien, Brun, Châtain, Noir, Roux... autant de couleurs avec autant de variations possibles. Pourtant, la coloration naturelle d'une chevelure n'est obtenue qu'à partir de seulement 2 mélanines (*annexe A.5.1*). C'est la variation de couleur de ces 2 mélanines ainsi que leurs concentrations qui permet une palette infinie allant du blond le plus clair au noir le plus profond.

La couleur d'un cheveu peut aussi changer en fonction de différents paramètres :

- **L'âge** : au cours de la vie, la couleur d'une chevelure d'une personne peut varier. L'évolution de la production de la mélanine dans la constitution d'un cheveu varie jusqu'à son arrêt total qui entraîne la non-coloration du cheveu qui pousse alors blanc (*annexe A.5.2*).
- **L'illumination** : la couleur d'une surface dépend de sa texture et de la réflexion de la lumière sur celle-ci [Longère 98]. En fonction de l'orientation de la source de lumière et de la "rugosité" de la surface, il apparaît des zones de reflets et d'ombres portées plus ou moins étendues. Un cheveu est considéré comme un cylindre semi-transparent [Marschner 03]. Une portion de la lumière se réfléchit directement sur la surface de la fibre (réflexion spéculaire de la couleur de la lumière), tandis qu'une autre pénètre à l'intérieur de la fibre où elle est absorbée, puis restituée en partie vers l'extérieur (réflexion diffuse colorée par les pigments de la fibre). On peut ainsi retrouver de nombreuses variations de couleurs simplement dues à l'illumination (*figure 5.1*). En général, nous pouvons décrire 3 zones : une zone claire de reflets spéculaires, une zone "normale", et une zone d'ombres. Cette approche a été notamment prise en compte par Chen *et al* dans leurs travaux de synthèse capillaire. [Chen 06].



FIGURE 5.1 – **Illumination d'une chevelure** : variations de couleur d'une chevelure en fonction de son illumination

- **L'environnement** : la chevelure subie des altérations qui peuvent être provoquées par une carence alimentaire, les intempéries, les agressions mécaniques (usage quotidien des sèche-cheveux, l'application de produits coiffants non adaptés à la texture des cheveux ou la mise en plis), par les rayons du soleil (destruction des écailles), par le stress, par la pollution de l'environnement et par l'eau.
- **L'utilisation de colorant** : la couleur de la chevelure peut être aussi changée artificiellement par l'utilisation de coloration permanente ou semi-permanente (*annexe A.5*). Une chevelure ne dépend plus seulement de sa coupe mais aussi de sa couleur. Ainsi de nouvelles tendances de coloration sont apparues :
 - *Coloration naturelle* : c'est l'application première des colorations afin de changer la couleur d'une chevelure tout en conservant des teintes naturelles.
 - *Coloration non naturelle* : dérive logique de la coloration capillaire, l'utilisation de teintes très différentes et non naturelles telles que le rouge, le bleu, le violet ...
 - *Coloration partielle* : c'est une nouvelle tendance, les colorations partielles de la chevelure soit en ne changeant que la couleur des pointes des cheveux, soit en faisant des coiffures à mèches, alternant des mèches de couleurs différentes.
 - *Décoloration* : la contrepartie de l'utilisation de teinture est la dégradation au cours du temps de la couleur de la chevelure : La teinture qui déteint au fil des lavages en particulier. La repousse des cheveux entraîne aussi une différence de couleur entre les racines et les pointes.

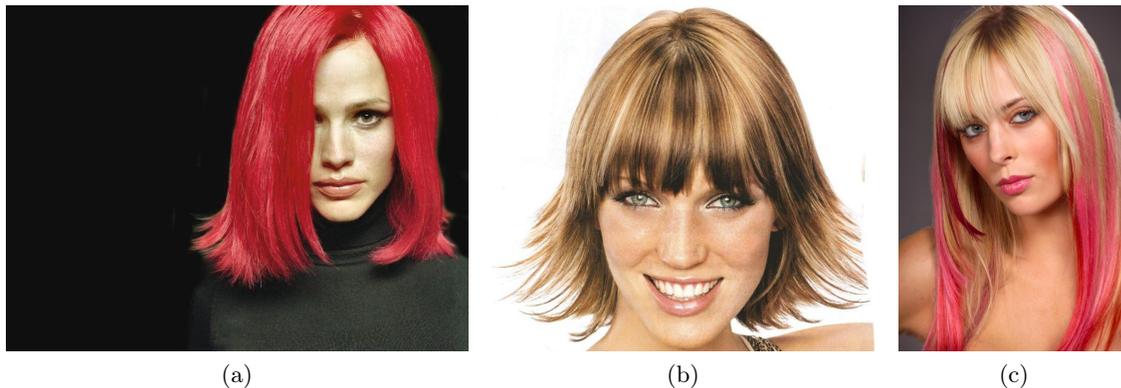


FIGURE 5.2 – **Coloration Non Naturelle** (a) Coloration rouge non naturelle, (b) Coiffure avec des mèches de différentes teintes, (c) Coloration non naturelle avec des mèches de différentes teintes

5.1.2 La segmentation couleur

Dans la section 2.2, nous avons présenté une revue non exhaustive de différentes techniques de segmentation. La segmentation couleur, bien qu'elle puisse être utilisée pour des méthodes par approche "*contours*", s'apparente plus aux méthodes par approches "*régions*" qui consistent à rechercher des ensembles de pixels qui possèdent des propriétés communes.

Nous pouvons regrouper ces méthodes en deux catégories :

- **La segmentation couleur non supervisée** : elle vise à séparer automatiquement l'image en régions naturelles, c'est-à-dire sans aucune connaissance préalable des classes. Ces méthodes considèrent une région comme un ensemble de pixels connexes dont les couleurs sont proches les unes des autres. Elles procèdent par un balayage de l'image afin de construire les régions et font intervenir l'information colorimétrique pour la prise de décision. Nous retrouvons les approches connues par division de régions, par croissance de régions, par fusion de régions et par division-fusion de régions. On peut citer par exemple les travaux de Trémeau *et al* [Trémeau 97] qui proposent un algorithme de segmentation qui combine une croissance de régions suivie d'un processus de fusion de régions.
- **La segmentation couleur supervisée** : elle divise l'image en différentes classes à partir d'information a priori qui est introduite sous la forme d'échantillons d'apprentissages. Cette méthode de classification ne prend pas en compte la disposition spatiale des pixels et ne considère que la distribution des couleurs dans l'espace de représentation utilisé. Elle identifie les classes de pixels en présence dans l'image et affecte à chaque pixel une étiquette indiquant la classe à laquelle il appartient. La formation des régions n'est obtenue qu'après analyse de la connexité des pixels dans l'image étiquetée.

Le but de notre analyse de la teinte capillaire est de créer la carte de localisation couleur de la chevelure pour la définition des marqueurs pour le traitement "*Matting*". Ainsi nous devons trouver les pixels de l'image correspondant à une classe couleur de la chevelure. Nous allons donc orienter notre approche par une segmentation couleur supervisée. Mais comme nous l'avons expliqué dans la section précédente, la définition d'un modèle de teinte pour la chevelure est

complexe. Nous allons donc modéliser l'information couleur à partir de l'étude de fenêtres échantillons. Afin de déterminer la position de ces zones, nous nous sommes intéressés à la segmentation du masque de peau afin de localiser le visage dans une image. La segmentation couleur de la peau est aussi une étape importante de notre algorithme. En effet même si nous ne connaissons pas la position exacte de la chevelure dans une image, nous pouvons admettre facilement qu'elle se trouve autour d'un visage. Contrairement à l'analyse capillaire, la détection et la segmentation de zones de peau a fait l'objet de nombreux travaux de recherche notamment par des approches de segmentation couleur supervisée [Jones 99] [Kakumanu 06].

5.1.3 Base de référence de masque de peau

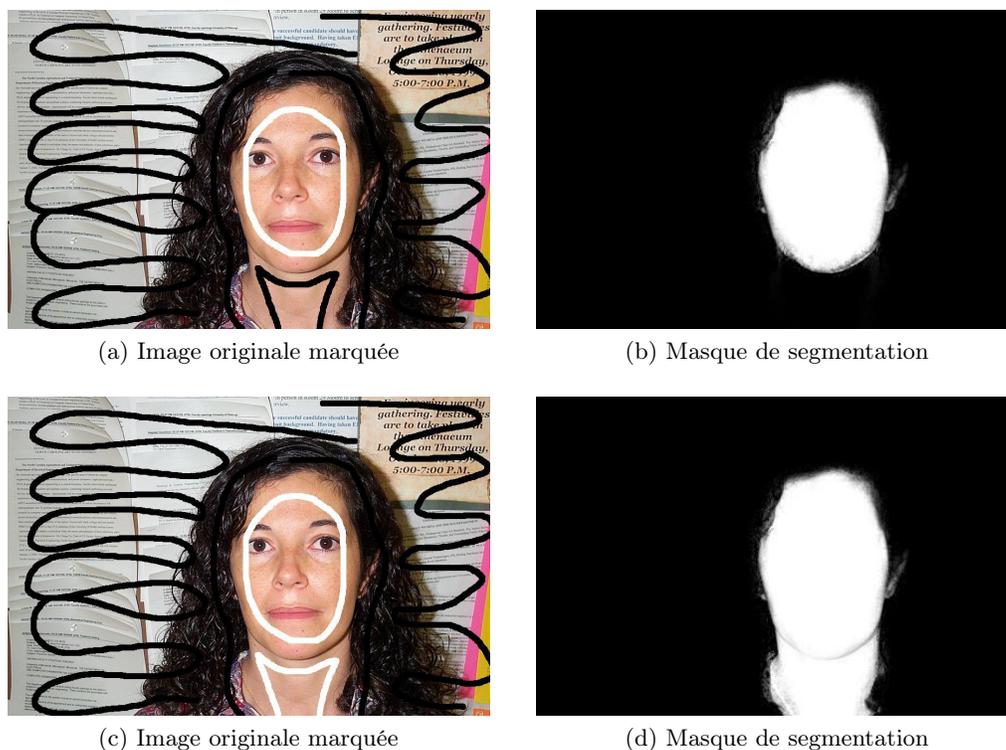


FIGURE 5.3 – **Base de Référence Visage** : 1^{ère} ligne : masque de peau du visage seulement, 2^{ème} ligne : masque de peau du visage et du cou

Pour évaluer nos résultats de segmentation du masque de peau, comme pour la base de référence capillaire, nous avons créé une base de référence semi-manuellement à l'aide du traitement "Matting" initialisé par des marqueurs placés manuellement. Nous avons créé la base de référence à partir des images de la base Caltech (section 3.4.3). Cependant, contrairement aux images testées pour la segmentation capillaire, nous avons pris indifféremment des modèles masculins et féminins. On peut aussi signaler la présence de photos avec des modèles féminins portant une frange sur le front. Nous avons délibérément conservé ces images dans la base afin de tester notre approche aux plus grands nombres de cas possibles.

Deux bases d'images ont été ainsi créées : une base contenant seulement le visage et une base dans laquelle le cou a été aussi pris en compte. La prise en compte du cou a deux objectifs. Le

premier est simplement pour augmenter la zone de peau connue afin de limiter le taux de fausse acceptation de nos résultats. La seconde est de voir si la segmentation couleur est assez efficace pour la segmentation du cou qui est souvent une zone sous-exposée à la lumière. Pouvoir créer une carte de localisation incluant le cou pourra être utile pour une meilleure prise en considération de chevelures longues.

5.2 La Segmentation Couleur supervisée

Dans un processus de segmentation couleur par apprentissage, il existe 2 étapes majeures :

- la phase d’entraînement qui se divise en 3 étapes : la collecte d’une base d’image de référence, le choix d’un espace couleur et l’apprentissage d’un classifieur couleur.
- la phase de détection qui consiste à classifier chaque pixel binairesment selon qu’il corresponde ou non au modèle couleur.

5.2.1 Le choix d’un espace couleur

Le choix de l’espace couleur adapté à une application de segmentation couleur est toujours complexe. En effet comme nous l’avons vu dans la section précédente, il en existe un grand nombre de représentations possédant toutes des avantages et des inconvénients, que ce soit dans l’information portée par chacune des composantes ou dans la dynamique de représentation de ces composantes. Ainsi le choix d’un espace couleur dépend de plusieurs paramètres : la nature des éléments à segmenter (forme simple géométriques ou objets complexes, forme de couleur unie ou de plusieurs composantes couleurs) et de la base d’image à traiter (image synthétique/naturelle, type d’illumination, résolution).

En fonction de ces paramètres, il est possible d’orienter son choix pour un espace plutôt qu’un autre. Néanmoins, on retrouve souvent une utilisation d’un grand nombre des espaces couleurs pour des applications du même type. Par exemple, bien que la littérature sur la segmentation capillaire soit très faible, on dénombre l’utilisation de pas moins de 5 espaces couleurs en quasiment autant de publications : l’espace RVB [Yacoob 06], [Lee 08], [Wang 09], l’espace Luv [Chen 06], l’espace La^*b^* [Lee 08], l’espace XYZ [Zhang 08] et l’espace HSV [Zhang 08].

En se basant sur les travaux de Elgammal *et al* [Elgammal 09], nous allons voir les avantages et inconvénients des espaces couleurs cités précédemment pour l’analyse couleur de la peau. Nous avons testé ces différents espaces pour différents types de couleur de peau (*figure 5.4*).

Le choix de l’espace couleur dépendra de la représentation du modèle couleur, de sa forme (compacte ou pas) et de son invariabilité (variation en fonction de la couleur/luminance) :

- **L’espace RVB** : De par sa représentation, l’espace RVB est le seul espace à ne pas séparer la luminance et la chrominance. Nous représentons la distribution couleur selon les composantes Rouge et Verte car ce sont les composantes auxquelles l’oeil est le plus sensible. On constate que les distributions couleurs pour les 3 types de peaux sont étendues et que pour un faible nombre de pixels dans la fenêtre échantillon déjà un grand nombre de valeurs sont présentes (*figures 5.4d 5.4e 5.4f*). Cette répartition peut s’expliquer par la forte corrélation entre les 3 composantes RVB qui implique qu’une faible variation

d'intensité d'illumination de la peau va affecter la représentation couleur selon les 3 composantes dans le cube RVB . De plus la localisation des distributions couleurs est très variable particulièrement selon la composante rouge.

Malgré cette limitation, l'espace RVB reste largement utilisé dans la segmentation de peau de par sa simplicité et des résultats satisfaisants qu'il permet d'obtenir [Kovac 03] [Jedynak 03] [Harville 05] [Shin 02].

- **L'espace $CbCr$** : Dans l'espace $YCbCr$, la description de la couleur de peau s'appuie seulement sur les 2 composantes de chrominance (Cb, Cr) et ainsi ne sera pas affectée par des variations d'illumination dans le plan Y . La distribution couleur est assez compacte pour les 3 types de couleurs de peau et on peut la représenter généralement par une ellipse dans le plan de chrominance. L'autre point intéressant à souligner est que la représentation pour les différents types de couleur de peau est très proche dans les plans de chrominance (*figures 5.4j 5.4k 5.4l*).

Cette distribution couleur peut être facilement modélisée par une distribution gaussienne. Ainsi ses propriétés d'invariance et sa simplicité de transformation à partir de l'espace RGB , font de l'espace $YCbCr$ le plus couramment utilisé dans la segmentation de la peau [Kovac 03] [Cho 07] [Chai 99] [Shin 02] [Torres 01] [Chai 98].

- **L'espace HSV** : Le principal intérêt de cet espace est la description intuitive de la couleur de peau selon les composantes de teinte H et de saturation S (*figures 5.4g 5.4h 5.4i*). Bien que les effets de variation d'illumination soient principalement portés par la composante de luminosité V , on constate une représentation assez étendue et une variation de la localisation de la distribution assez similaire à l'espace RVB . C'est un espace couleur qui demeure peu utilisé [Shin 02] [Sobottka 96].
- **L'espace La^*b^*** : Comme pour l'espace $YCbCr$, la description de la couleur de peau ne se base que sur les composantes de chrominances. D'ailleurs la représentation de la couleur de peau est très similaire à l'espace $YcbCr$ et peut être définies par des ellipses compactes (*figures 5.4m, 5.4n, 5.4o*). Cependant, bien que cet espace apporte une représentation intéressante, sa difficulté d'implémentation et son temps de traitement assez long comparé à l'espace $YcbCr$ font qu'il est finalement peu utilisé dans le domaine [Shin 02].

De par ces 2 spécificités, nous constatons que l'espace $YCbCr$ semble le plus approprié des espaces couleurs étudiés. Contrairement aux espaces RGB et HSV , la densité de la distribution de couleur de peau est plus compacte. On peut aussi observer que la localisation du modèle couleur de peau est plus stable pour les 3 types couleur de peau avec une localisation très proche, ce qui montre que l'espace $YCbCr$ est particulièrement adapté et stable pour des variations de teintes de peau dues à des origines ethniques mais aussi pour des variations des conditions d'illuminations.

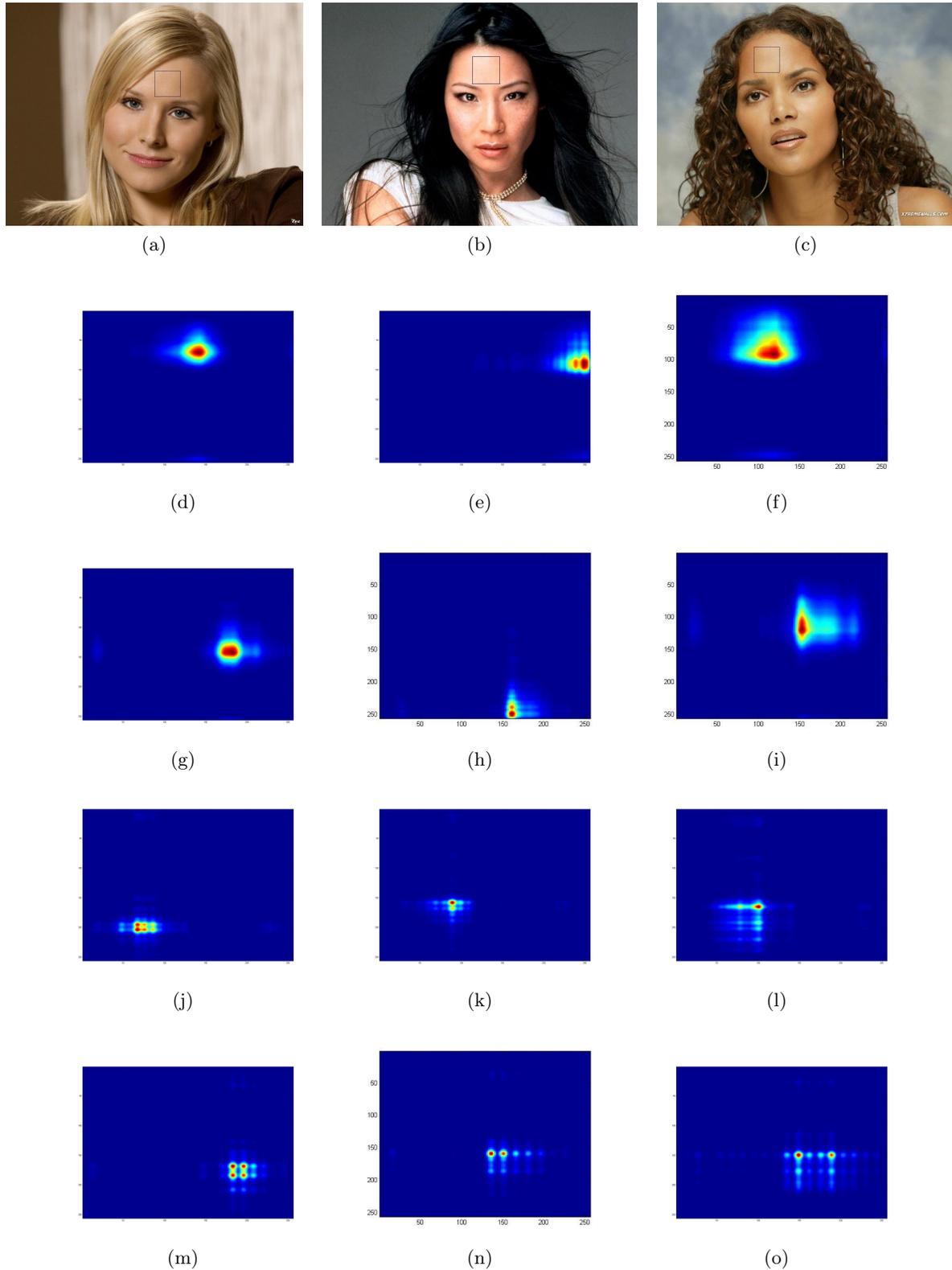


FIGURE 5.4 – Représentation de la couleur de peau d'un patch de front pour différents espaces couleurs Représentation de la couleur de peau pour l'espace RVB (2^{ème} ligne), l'espace HSV (3^{ème} ligne), l'espace $YCbCr$ (4^{ème} ligne), l'espace La^*b^* (5^{ème} ligne)

5.2.2 La collecte d'une base d'image de référence couleur

Afin de modéliser au mieux un modèle couleur, généralement une base d'échantillons couleurs représentatifs est utilisée. Cette base doit contenir un grand nombre d'échantillons pris sous différentes conditions d'illumination et/ou de couleur (de peau, de chevelure) afin d'être le plus robuste aux variations possibles d'une image à une autre.

Dans leurs applications de segmentation capillaire, Zhang *et al* [Zhang 08], Lee *et al* [Lee 08] et Wang *et al* [Wang 09] modélisent ainsi des modèles couleurs de la chevelure à partir d'échantillons segmentés manuellement. Cette approche, en plus d'être fastidieuse, nécessite en particulier pour les modèles de chevelure des échantillons qui sont assez délicats à définir. En effet comme nous avons vu dans la section 5.1.1, la teinte peut être très variable au sein même d'une seule chevelure. Les bases d'entraînements devront donc pouvoir prendre en compte les variations d'illuminations (reflets spéculaires et zones d'ombres) ainsi que la variété de couleurs (blond, brun, noir...). Il faut ainsi pour chaque type de couleur définir des échantillons assez distincts pour ne pas inclure plusieurs teintes possibles. Par exemple pour une chevelure du type blonde, nous pouvons trouver jusqu'à trois classes possibles : Cheveux blanc pour les zones de reflets spéculaires, cheveux châtain pour les zones d'ombres et blond pour le reste de la chevelure sans parler de la possibilité que les racines appartiennent à une quatrième classe noir.

Dans [Yacoob 06], Yacoob *et al* définissent leurs modèles couleur directement à partir de l'image à traiter. En considérant des données topologiques du visage, ils définissent la position de fenêtres échantillons pour le calcul de leurs modèles. Nous pouvons aussi remarquer que malgré la définition de leurs modèles couleur à partir d'un apprentissage, Lee *et al* [Lee 08] et Wang *et al* [Wang 09] les mettent à jour aussi à partir de l'étude de l'image à traiter. Donc l'information locale de couleur paraît nécessaire pour la définition d'un modèle couleur optimal. Nous avons testé pour savoir si cette information était suffisante en comparant une approche de segmentation couleur dont le modèle est défini par apprentissage global ou local.

Pour la segmentation de peau, plusieurs zones références du visage sont possibles, les 2 zones les plus fréquemment utilisées sont :

- **Le front** : la partie centrale du front (*figure 5.5a*).
- **Les joues** : les 2 zones symétriques positionnées de chaque côté du nez sous les yeux (*figure 5.5b*).

En général, on recherche au maximum des zones localement homogènes lors de la définition des d'échantillons couleurs. A l'opposé, afin de prendre en considération plus de détails du visage et de pouvoir prendre en compte aussi l'orientation du visage ou la direction de la lumière, nous avons décidé de prendre une fenêtre échantillon autour du nez.

- **Le nez** : la zone centrale du visage située sous les yeux (*figure 5.5c*).

Avec cette fenêtre échantillon autour du nez, nous prenons en considération les reflets spéculaires sur la peau grâce à la pointe du nez, ainsi que les ombres portées sur le visage qui apparaissent aussi localement autour du nez. En prenant une fenêtre échantillon autour du nez plutôt que sur le front, nous évitons aussi les risques d'occultation possible par des cheveux.

Pour initialiser notre algorithme de détection de la peau nous avons d'abord besoin de lo-

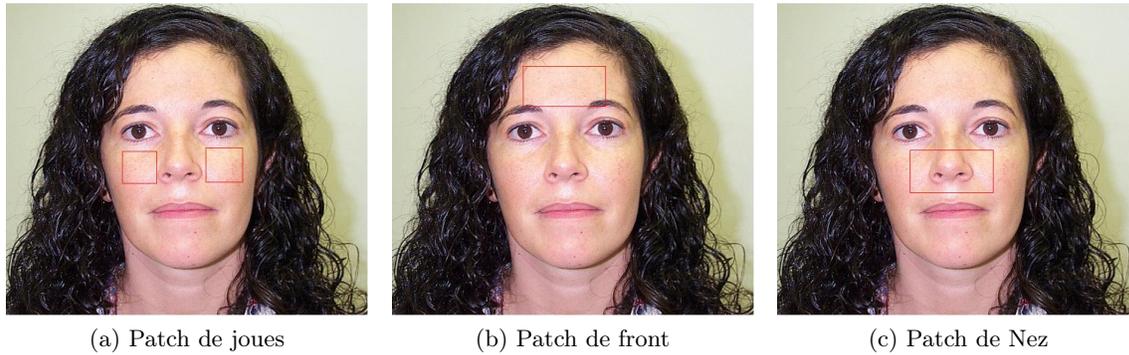


FIGURE 5.5 – Localisation des zones de patch de peau

caliser le visage dans l'image. Pour cela nous utilisons l'algorithme classique de détection de visage par modèle d'apparence : l'algorithme Adaboost en cascade de Viola & Jones [Viola 01] [Viola 04]. La détection du visage permet la définition d'une boîte englobante encadrant le visage. Nous couplons la détection du visage à une détection des yeux qui permet de définir les positions des centres de l'oeil gauche et droit ainsi qu'une boîte englobante pour chaque oeil.

Ainsi nous pouvons définir automatiquement une localisation des patches échantillons à partir des données morphologiques du visage que nous pouvons adapter en taille en fonction de la taille et du placement des yeux à l'intérieur du visage. Nous définissons la position de nos fenêtres échantillons selon les hypothèses suivantes :

- **Patch du Front :**
 - Hauteur Patch Front = Demi Distance entre les 2 yeux
 - Largeur Patch Front = Distance entre les 2 yeux
 - Localisation Patch Front = Quart de Distance entre les 2 yeux au-dessus les centres des yeux
- **Patch des Joues :**
 - Hauteur Patch Joue = Demi Distance entre les 2 yeux
 - Largeur Patch Joue = Demi Distance entre les 2 yeux
 - Localisation Patch Joue = Quart de Distance entre les 2 yeux, centré sur chaque oeil
- **Patch du Nez :**
 - Hauteur Patch Nez = Demi Distance entre les 2 yeux
 - Largeur Patch Nez = Distance entre les 2 yeux
 - Localisation Patch Front = Quart de Distance entre les 2 yeux au-dessous les centres des yeux

Nous avons comparé les résultats de segmentation du masque de peau pour définir quelle approche de segmentation couleur supervisée nous allons utiliser. Nous évaluons les différentes méthodes de définition des fenêtres échantillons pour 300 images de la base Caltech. Les résultats sont calculés sur les masques de peau obtenus après segmentation pour chaque position de la fenêtre échantillon pour un modèle couleur locale et un modèle couleur global. Avec cette expérience, nous avons testé 2 paramètres :

- *La localisation du patch* : Définir quelle est des 3 zones échantillons, celle qui est la plus stable pour la définition d'un modèle couleur de peau.
- *La définition d'un modèle locale ou globale* : Comparer les segmentations obtenues avec

un modèle couleur local calculé à partir d'une fenêtre échantillon unique avec un modèle couleur global calculé à partir d'une base d'apprentissage comprenant l'ensemble des échantillons pour une position donnée sur le visage (front, joues ou nez)

	Taux de Faux Rejet	Taux de Fausse Acceptance
modèle local Front	7.9%	82.4%
modèle global Front	5%	139.5%
modèle local Joue	8.2%	68.3%
modèle global Joue	5.36%	75.8%
modèle local Nez	8.3%	62.3%
modèle global Nez	5.59%	66.5%

TABLE 5.1 – **Évaluation de la définition du modèle couleur** : évaluation réalisée sur 100 images de la base Caltech

Nous observons que les taux de faux rejet (T.F.R.) sont très proches avec des résultats d'environ 5% pour les modèles globaux obtenus pour les 3 bases de références de 100 échantillons de front, des joues ou du nez. Les résultats pour les modèles locaux obtenus à partir de seulement un patch couleur sont très bas aussi avec seulement 7.9% pour le patch de front, 8.2% pour le patch de joues et 8.3% pour le patch de nez. Ces résultats sont intéressants pour 2 points, le premier montre que le choix d'un patch autour du nez apporte de bons résultats similaires aux deux autres patches. De plus, nous voyons que la définition du modèle couleur à partir d'un seul patch échantillon permet une segmentation donnant de bons résultats avec très peu de perte par rapport aux bases respectives de références.

Pour le taux de fausse acceptance (T.F.A.), les résultats sont plus contrastés. Les patches pris autour du nez donnent moins de sur segmentation avec seulement entre 62.3% et 66.5% , entre 86.3% et 97.8% pour les patches des joues, 82.3% pour le patch unique du front et jusqu'à 139.5% pour la base de référence front. Ces résultats s'expliquent notamment par la présence du modèle féminin portant une frange mais on peut aussi noter la présence de certains pixels de sourcils présents dans les patches de front pour les modèles masculins.

Les résultats obtenus permettent de valider notre hypothèse de travail sur la segmentation couleur de la peau à partir d'un unique patch de couleur défini autour de la région du nez.

Une amélioration possible pour optimiser la définition du modèle couleur d'un patch de front aurait été de prendre en compte la carte de fréquence définie précédemment afin de ne conserver que les pixels de peau. En effet en ne conservant que les pixels couleur hors de la carte fréquentielle, nous enlèverons les possibles pixels de cheveux et de sourcils présents dans la fenêtre échantillon pour ainsi obtenir une modélisation plus correcte du modèle couleur. Cependant, l'utilisation de la fenêtre échantillon autour du nez permet de simplifier la segmentation et permet de s'affranchir de la mauvaise définition de la carte fréquentielle.

5.3 Segmentation couleur de la peau

De nombreuses techniques de segmentation ou de classification de pixel de peau ont été développées dans la littérature. Typiquement un détecteur de peau est un classifieur une-classe qui

définit un espace de décision binaire dans l'espace de représentation couleur. Ainsi chaque pixel est classifié en tant que pixel de peau ou pixel non peau. La définition d'un classifieur est dépendante de la représentation de l'espace de décision dans l'espace de couleur. Plus la représentation du modèle de couleur de peau sera compacte, plus le classifieur aura une représentation simple et permettra une prise de décision facile.

5.3.1 Classifieur couleur par valeurs fixes

Une première approche basique est de déterminer la dynamique couleur des pixels de peau. Pour des conditions d'acquisition bien contrôlée, on peut définir des seuils pour chaque composante couleur.

– *Dans l'espace RVB :*

Peer *et al* [Kovac 03] ont défini des valeurs seuils de la peau dans l'espace RVB en prenant en compte une information d'orientation et d'uniformité de l'illumination de la personne.

Pour un sujet sous une illumination uniforme, les pixels de peau sont définis selon les équations suivantes :

$$\begin{aligned} R > 95 \quad \text{et} \quad G > 40 \quad \text{et} \quad B > 20 \\ \max\{R, G, B, \} - \min\{R, G, B, \} > 15 \\ |R - G| > 15 \\ R > G \quad \text{et} \quad R > B \end{aligned} \tag{5.1}$$

Pour un sujet dont l'illumination est latérale, les pixels de peau sont définis selon les équations suivantes :

$$\begin{aligned} R > 220 \quad \text{et} \quad G > 210 \quad \text{et} \quad B > 170 \\ \max\{R, G, B, \} - \min\{R, G, B, \} > 15 \\ |R - G| \leq 15 \\ R > G \quad \text{et} \quad R > B \end{aligned} \tag{5.2}$$

– *Dans l'espace YCbCr :*

Une approche similaire a été utilisé par Chai *et al* [Chai 99] dans l'espace YCbCr. Les composantes de chrominances n'étant pas autant corrélées que pour l'espace RVB, un simple encadrement des valeurs possibles de chaque composante pour les pixels de peau est suffisant.

$$\begin{aligned} 77 &\leq Cb \leq 127 \\ 133 &\leq Cr \leq 173 \end{aligned} \tag{5.3}$$

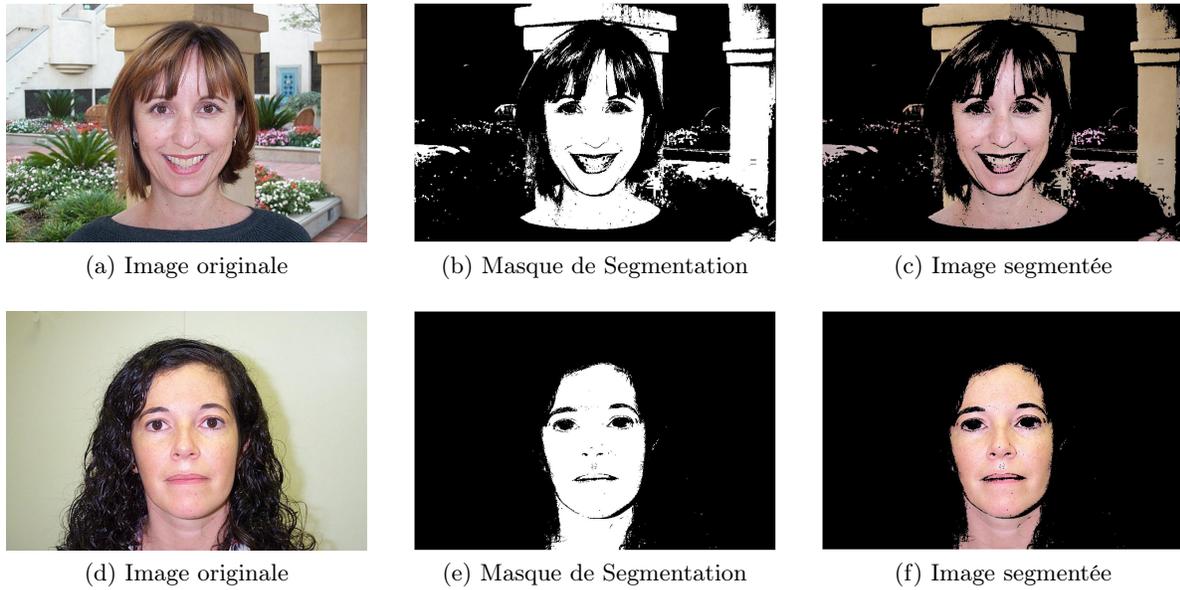


FIGURE 5.6 – Segmentation couleur par valeur fixes dans l'espace RVB

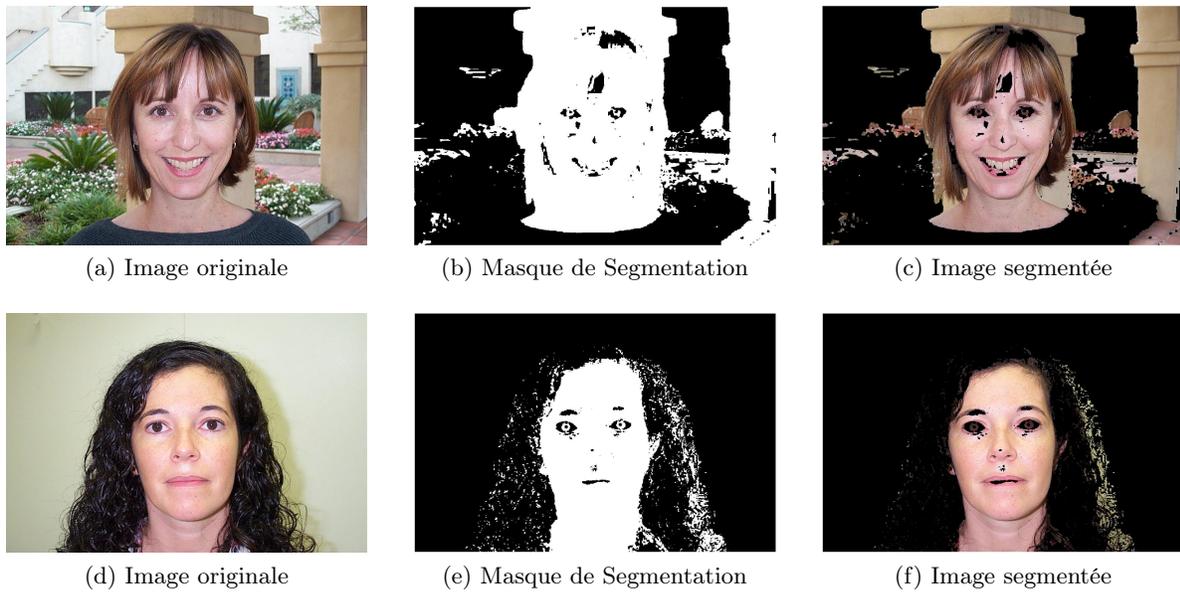


FIGURE 5.7 – Segmentation couleur par valeur fixes dans l'espace $YCbCr$

5.3.2 Classifieur Couleur par Minimisation de la distorsion de couleur

Nous présentons ici le classifieur couleur utilisé par Yacoob *et al* [Yacoob 06] basé sur les travaux Horprasert [Horprasert 99]. Il appuie son modèle couleur sur la modélisation synthétique d'un cheveu selon Kajiya *et al* [Kajiya 89] qui utilisent une modélisation par une surface Lambertienne afin de prendre en compte la diffusion de la lumière dans un cheveu. Ainsi pour un cheveu considérée comme une surface Lambertienne, une couleur dépend de sa luminance et de réflectance spectrale, paramètres qui sont les hypothèses de recherche de Horprasert *et al* pour la modélisation de leur modèle couleur.

Dans la figure 5.8 la ligne OE_i représente la ligne de chromaticité voulue dans l'espace couleur RGB, avec $E_i = (E_r(i), E_g(i), E_b(i))$ représentant la couleur modélisée pour le pixel i et nous définissons $I_i = (I_r(i), I_g(i), I_b(i))$ la couleur du pixel courant.

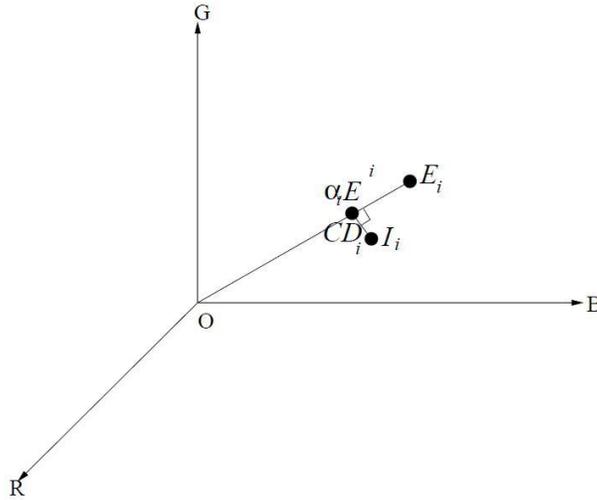


FIGURE 5.8 – **Segmentation couleur par Minimisation de la distorsion de couleur** : Représentation du Modèle couleur dans l'espace RGB

La *distorsion de luminance* entre les points E_i et I_i correspond à la différence de la projection du point I_i sur la ligne OE_i est modélisée par l'équation :

$$f(\alpha_i) = (I_i - \alpha_i E_i)^2 \quad (5.4)$$

avec α_i représentant la différence de luminance, $\alpha > 1$ si le pixel courant est plus clair que le modèle et $\alpha < 1$ si le pixel est plus sombre.

La *distorsion couleur* CD_i mesurant la distance orthogonal entre le pixel courant et la ligne de chromaticité est définie selon l'équation :

$$CD_i = \| I_i - \alpha_i E_i \| \quad (5.5)$$

Le modèle couleur de référence est calculé afin de définir un 4-tuple $\langle E_s, \sigma_s, a_s, b_s \rangle$ avec $E_s = [\mu_r(s), \mu_g(s), \mu_b(s)]$ la valeur moyenne pour chaque composante couleur, $\sigma_s = [\sigma_r(s), \sigma_g(s), \sigma_b(s)]$ l'écart type, et 2 termes de normalisation $a_s = \sqrt{\frac{\sum_{i=0}^N (\alpha_i - 1)^2}{N}}$ qui représente la variation de la

distorsion de luminance et $b_s = \sqrt{\frac{\sum_{i=0}^N (CD_i - 1)^2}{N}}$ qui représente la variation de la distorsion de chromaticité.

Ainsi les équations de la distorsion de Luminance (5.4) et de la distorsion de Chromaticité (5.5) sont redéfinies selon les équations :

$$\alpha_i = \frac{\left(\frac{I_r(i)\mu_r(i)}{\sigma_r(i)} + \frac{I_g(i)\mu_g(i)}{\sigma_g(i)} + \frac{I_b(i)\mu_b(i)}{\sigma_b(i)} \right)}{\left(\frac{\mu_r(i)}{\sigma_r(i)} \right)^2 + \left(\frac{\mu_g(i)}{\sigma_g(i)} \right)^2 + \left(\frac{\mu_b(i)}{\sigma_b(i)} \right)^2} \quad (5.6)$$

$$CD_i = \sqrt{\left(\frac{I_r(i) - \alpha_i \mu_r(i)}{\sigma_r} \right)^2 + \left(\frac{I_g(i) - \alpha_i \mu_g(i)}{\sigma_g} \right)^2 + \left(\frac{I_b(i) - \alpha_i \mu_b(i)}{\sigma_b} \right)^2}. \quad (5.7)$$

Au final, un pixel est classifié comme un pixel correspondant au modèle couleur si la distorsion de chromaticité CD_i est inférieure à un seuil prédéfini et que la distorsion de luminance α_i est comprise entre 2 seuils déterminés en fonction du taux de détection voulu dans la base d'apprentissage du modèle couleur.



(a) Image originale



(b) Masque de Segmentation



(c) Image segmentée



(d) Image originale



(e) Masque de Segmentation



(f) Image segmentée

FIGURE 5.9 – Segmentation couleur par Minimisation de la distorsion de couleur

5.3.3 Classifieur Couleur Probabiliste

Une approche possible est de considérer la segmentation couleur comme un problème probabiliste [Zhang 08], [Lee 08] [Wang 09]. Etant donné un pixel de couleur c , quelle est la probabilité que ce pixel soit un pixel correspondant à notre modèle couleur $P(\text{modele}|c)$. Une fois cette probabilité calculée, la segmentation est effectuée en donnant une étiquette aux pixels en fonction d'un seuil défini. Le calcul de telles probabilités pour chaque couleur est évidemment impossible (il faudrait calculer 256^3 probabilité dans l'espace RVB par exemple).

En utilisant la relation de Bayes le problème peut se traduire par l'équation :

$$P(\text{modele}|c) = \frac{P(c|\text{modele})P(\text{modele})}{P(c|\text{modele})P(\text{modele}) + P(c|\text{modele}^C)P(\text{modele}^C)} \quad (5.8)$$

Ainsi selon le modèle de Bayes la probabilité a posteriori $P(\text{modele}|c)$ pour un pixel d'appartenir au modèle couleur en fonction de sa couleur est fonction de la fonction de vraisemblance du modèle couleur $P(c|\text{modele})$ et de la probabilité a priori $P(\text{modele})$ qui estime la probabilité qu'un pixel appartienne au modèle couleur. Le problème de classification se ramène au final au calcul du terme de vraisemblance $P(c|\text{modele})$.

Différentes approches ont été introduites pour calculer cette fonction de densité de probabilité :

- *Modèle des Histogrammes* [Jones 04] : Un histogramme est calculé à partir de la distribution couleur du modèle. Les valeurs de l'histogramme sont ensuite converties en une distribution de probabilité discrète selon l'équation :

$$P(c|\text{modele}) = \frac{s[c]}{T_s} \quad (5.9)$$

où $s[c]$ est la valeur de l'histogramme pour la couleur c et T_s le nombre total de valeurs de l'histogramme.

- *Modèle simple Gaussienne* : On peut représenter la distribution du modèle couleur selon une distribution d'une variable aléatoire D-dimensionnelle X définie par l'équation :

$$P(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{D}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma (x - \mu)\right) \quad (5.10)$$

où $\mu = \frac{1}{n} \sum_{i=1}^n c_i$ est la couleur moyenne et $\Sigma = \frac{1}{n-1} \sum_{i=1}^n (c_i - \mu)(c_i - \mu)^T$ la matrice de covariance de la variable X.

- *Modèle par Mélange de Gaussienne (GMM)* [Yang 99] : Bien que le modèle par simple gaussienne donne de bons résultats et permet la distinction de différentes classes, la distribution d'un modèle couleur n'est pas forcément unimodale. Une meilleure approximation peut être alors obtenue si on considère la distribution comme la somme de multiples sources gaussiennes indépendantes.

Dans ce cas, la fonction de répartition peut être estimée en utilisant un nombre fini de gaussiennes mélangées. La fonction de densité de probabilité de GMM peut être définie

comme une somme pondérée de gaussiennes selon l'équation :

$$P(x; \theta) = \sum_1^N \alpha_i G(x; \mu_i, \sigma_i) \quad (5.11)$$

en notant le vecteur paramètre :

$$\theta \{ \alpha_i, \mu_i, \sigma_i \} \quad \text{pour } i = 1..N \quad \text{où } \alpha_i \text{ est le terme de mélange} \quad (5.12)$$

L'évaluation du modèle de GMM est effectuée suivant l'algorithme EM (Expectation-Maximization) [Dempster 77] qui est basé sur une optimisation itérative du vecteur paramètre.

Rapidement, l'utilisation d'un mélange de gaussiennes pour la représentation de notre modèle couleur s'est avérée non adaptée et la représentation selon le modèle d'une simple gaussienne largement suffisant. En effet, comme nous considérons l'analyse à partir seulement d'une fenêtre échantillon, elle contient un faible nombre de pixels. Le calcul du mélange de gaussiennes optimisées par l'algorithme EM n'était pas performant, d'autant plus que la définition du nombre de gaussiennes possibles n'était pas fixe d'une fenêtre échantillon à l'autre. D'autre part la fenêtre échantillon étant assez bien localisée, la répartition du modèle couleur est assez compacte pour la définition d'une simple gaussienne englobante.

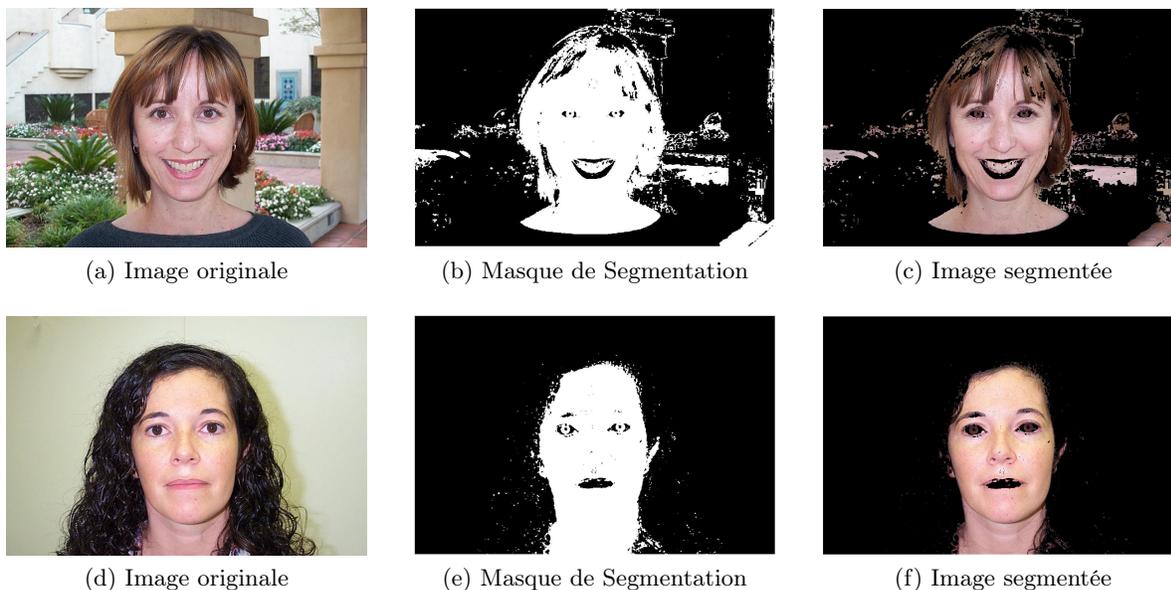


FIGURE 5.10 – Segmentation couleur par Simple Gaussienne 2D

À partir du patch échantillon couleur dans l'espace $YCbCr$, nous représentons sa distribution couleur sur laquelle nous effectuons un filtrage passe-bas gaussien afin d'uniformiser la distribution qui a été obtenue avec un faible nombre de pixels. Comme nous travaillons sur un classifieur couleur binaire, le pixel testé appartient ou n'appartient pas au modèle couleur, nous simplifions la définition de notre classifieur par approche probabiliste en définissant la prise de décision selon le modèle $P(c|\text{modele}) > \text{seuil}$. Nous fixons le seuil selon l'écart type de la distribution Gaussienne.

Nous faisons une analogie entre la distribution gaussienne du modèle et une distribution de probabilité. Plus un pixel couleur est proche de la moyenne de la distribution couleur plus sa probabilité d'être un pixel visage est grande. Inversement plus un pixel est loin de la moyenne de la distribution couleur plus sa probabilité d'être un pixel non visage est grande. Cette représentation va nous permettre de définir plusieurs degrés de probabilité d'appartenance au modèle couleur qui va nous être utile pour notre traitement de fusion de données par fonction de croyances (section 6.3.2.3).

Pour simplifier les calculs, nous avons mis en place une segmentation couleur par gaussienne une dimension définie pour chaque plan de chrominance Cb et Cr. Pour chaque composante couleur la moyenne $\mu_{(Cb,Cr)}$ et l'écart type $std_{(Cb,Cr)}$ de la distribution couleur sont calculés et un masque de segmentation est défini selon l'équation :

$$\text{Masque couleur}_{(Cb,Cr)}(x,y) \begin{cases} = 1 & \text{si } \begin{cases} Imf(x,y) > \mu_{(Cb,Cr)} + std_{(Cb,Cr)} \\ Imf(x,y) < \mu_{(Cb,Cr)} - std_{(Cb,Cr)} \end{cases} \\ = 0 & \text{sinon} \end{cases} \quad (5.13)$$

Le masque de peau final est obtenu par l'opération simple du ET logique.

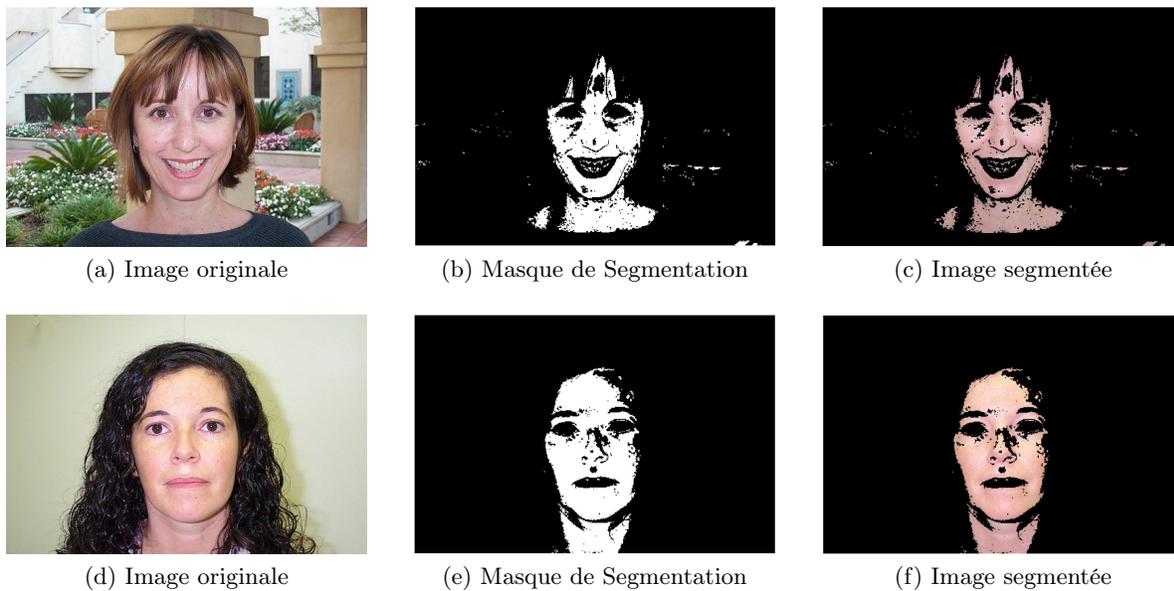


FIGURE 5.11 – Segmentation couleur par Simple Gaussienne 1D : masque de peau final obtenu par l'opération simple du ET logique

5.3.4 Classifieur Couleur de Peau

Pour choisir le classifieur couleur que nous allons utiliser nous avons comparé les différentes méthodes de segmentation couleur de la peau que nous avons étudié. Nous les évaluons pour 300 images de la base Caltech. Les résultats sont comparés entre les masques de peau obtenus après segmentation pour chaque classifieur couleur et les deux bases de références. Pour les classifieurs de minimisation de la distorsion couleur, par modélisation par simple gaussienne 2D et par simple gaussienne 1D, le modèle couleur est calculé automatiquement à partir de la fenêtre échantillon autour du nez.

	Taux Faux Rejet Visage	Taux Fausse Acceptance Visage
RGB Fixe	22.8%	43.2%
YCbCr Fixe	21.3%	152%
Distorsion Couleur	65%	60%
Gaussienne	8.3%	62.3%
Double Gaussienne	23.4%	29.7%

TABLE 5.2 – Évaluation de la segmentation couleur de pixels de peau du visage : Résultats obtenues pour la segmentation de 200 images de la base Caltech

	Taux Faux Rejet (<i>Visage & Cou</i>)	Taux Faux Acceptance (<i>Visage & Cou</i>)
RGB Fixe	23.8%	15.7%
YCbCr Fixe	21.6%	93%
Distorsion Couleur	70%	48.3%
Gaussienne	11.2%	45.3%
Double Gaussienne	28.6%	13%

TABLE 5.3 – Évaluation de la segmentation couleur de pixels de peau du visage et du cou : Résultats obtenues pour la segmentation de 200 images de la base Caltech

Pour les classifieurs couleurs par valeurs fixe, nous obtenons de bons résultats avec environ 20% de faux rejet et des taux de fausse acceptance plus important entre 15% et 40% pour la segmentation RGB et jusqu'à plus de 150% pour la segmentation Ycbr. Paradoxalement c'est avec le seuillage par valeur fixe dans l'espace RVB que nous avons de meilleurs résultats comparés à l'espace YCbCr. Même si l'ensemble des résultats sont satisfaisant la proportion de pixels sur-segmentés peut vraiment atteindre une grande partie de l'image.

Les résultats pour le classifieur par minimisation de la distorsion couleur montre que cette approche n'est pas adaptée pour la segmentation couleur de la peau avec près de 70% de taux de faux rejet. Cela implique que plus de la moitié des pixels de peau ne sont pas correctement segmentés. On constate un faible taux de fausse acceptance entre 50% et 60%. Ces résultats peuvent s'expliquer par la faible dynamique du classifieur et par une sous-segmentation générale de l'image. De plus ce classifieur possède des limites en particulier dues à l'illumination de l'image car c'est une approche basée sur une segmentation dans l'espace RVB, les mauvaises projections de certains pixels saturés entraînent une sur-segmentation des zones claires ou des zones sombres.

Les meilleurs résultats de segmentation de la peau sont obtenus avec le classifieur simple

gaussien 2D de l'espace YcbCr avec moins de 10% de taux de faux rejets. Néanmoins, même si la peau est particulièrement bien segmentée, le taux de fausse acceptation reste relativement élevé avec entre 45% et 60%. Le modèle par gaussienne 1D donne des résultats un peu moins bons avec seulement environ 25% de taux de faux refus. Ces résultats sont plus élevés mais restent dans l'ordre de grandeur des classifieurs par valeur fixe. Par contre, ce classifieur est vraiment plus performant car il limite la sur-segmentation avec un taux de fausse acceptation relativement bas entre 15% et 30%.

La différence entre les 2 classifieurs par gaussienne peut s'expliquer par plusieurs raisons. En considérant une gaussienne 1D, nous réduisons la dynamique pour chaque composante et nous diminuons la performance du classifieur ce qui entraîne une sous-segmentation de l'image. D'un autre côté, nous limitons la sur-segmentation de l'image avec un faible taux de faux rejet. Nous sommes aussi plus stables par rapport à la composante (128, 128) représentant l'axe des niveaux de gris dans le plan CbCr.

De plus dans l'optique de notre traitement, une sous-segmentation est préférée à la sur-segmentation. En effet notre analyse couleur est une initialisation au placement des marqueurs pour initialiser le traitement "*Matting*". Ainsi le but est d'identifier au mieux les pixels de chevelure potentielle, la segmentation étant assurée par l'estimation de l'alpha matte pendant le traitement "*Matting*". Ainsi nous choisirons pour la suite de notre étude, le classifieur couleur modélisé par une gaussienne 1D.

5.4 Segmentation couleur de la chevelure

Nous avons vu que la segmentation de peau est déjà un problème difficile pourtant la peau est une surface homogène et dont les composantes de chrominance sont assez stables quel que soit le type de couleur de peau (*figure 5.4*). En comparaison la segmentation couleur des cheveux est bien plus difficile car une chevelure est moins homogène et peut prendre une grande palette de teintes. Nous allons appliquer la méthode introduite pour la segmentation de peau et voir les résultats obtenus pour la segmentation capillaire.

Nous avons vu que l'utilisation d'un patch échantillon pour définir un modèle couleur donnait de bons résultats pour la segmentation de peau, nous baserons notre approche couleur de la chevelure sur le même principe. L'échantillon de couleur va permettre de définir un modèle adapté à la chevelure étudiée. La prise en compte de plusieurs patches de couleur de peau a été envisagée en se basant sur les travaux de Yacoob *et al* [Yacoob 06] qui définissent 3 fenêtres échantillons. Plusieurs paramètres montrent la limite d'une telle définition : comme nous l'avons défini en début de chapitre, une chevelure peut prendre plusieurs teintes au sein d'une même chevelure, ainsi il existe une probabilité de segmenter une chevelure appartenant à plusieurs modèles couleurs.

Dans [Yacoob 06], une différence de couleur est calculée entre les 3 modèles couleurs définis, si la distance entre les couleurs moyennes de chaque patch est trop grande alors seul le modèle du haut du visage est considéré comme modèle de référence. Cette approche est limitée et ne permet pas de prendre en compte de nombreuses coupes possibles dont notamment les personnes chauves ou du moins ayant une calvitie au sommet du crâne.

Comme nous limitons notre étude à des modèles féminins avec une chevelure, nous admettons dans une première approche que nous pouvons prendre en référence qu'un patch au sommet du

crâne. Comme pour le modèle couleur de peau nous définissons automatiquement la localisation du patch échantillon à partir des données morphologiques du visage. Nous définissons la position de notre fenêtre échantillon de chevelure selon les hypothèses suivantes :

– **Patch de Chevelure :**

- Hauteur Patch Cheveux = Distance entre les 2 yeux
- Largeur Patch Cheveux = 2 * Distance entre les 2 yeux
- Localisation Patch Cheveux = Demi Distance entre les 2 yeux au-dessus les centres des yeux

Nous optimisons la définition du modèle couleur de chevelure en ajoutant l'information de la carte de localisation fréquentielle. La carte fréquentielle représente la localisation la plus probable des zones fréquentielles propres à une chevelure. En ne conservant que les pixels de la fenêtre échantillon contenus dans la carte fréquentielle, nous affinons la modélisation du modèle couleur en ne prenant pas en compte les potentiels pixels de peau qui correspondent à des zones homogènes a priori. Nous aurions pu aussi utiliser la carte de localisation du visage obtenu précédemment mais sa définition n'était pas suffisamment précise pour éliminer tous les pixels de peau particulièrement autour des racines des cheveux.

Une fois le modèle couleur gaussien 1D calculé, nous segmentons l'image selon le classifieur couleur défini pour la segmentation du masque de couleur de peau par l'équation (5.13). Nous obtenons de bons résultats (*figure 5.12*) avec dans l'ensemble une bonne localisation d'une partie de la chevelure. Nous avons aussi une part importante de sur-segmentation de pixels de fond ou du visage en particulier pour les coiffures foncées dont la couleur sombre est proche de la couleur des zones d'ombres de l'image. Nous n'arrivons pas à segmenter généralement l'ensemble de la chevelure car comme nous l'avons déjà noté, la chevelure peut prendre plusieurs teintes ou du moins une grande variation de teinte. Notre segmentation couleur n'est pas optimale et pourrait être encore améliorée mais elle est néanmoins largement suffisante pour notre approche de segmentation capillaire par l'algorithme de "*Matting*". En effet notre carte de localisation de la chevelure en couleur doit représenter les pixels qui ont le plus de probabilité d'appartenir à la chevelure pour permettre la meilleure définition des marqueurs cheveux H_m et non cheveux \bar{H}_m

5.5 Conclusion

Dans ce chapitre, nous avons introduit nos travaux sur l'analyse couleur de la chevelure. Nous avons montré dans un premier temps les caractéristiques de la représentation de la chevelure dans une image. Puis dans un second temps, nous avons présenté en détail la segmentation couleur par approche supervisée et la définition d'un classifieur couleur à partir de modèles couleurs. Nous avons défini notre espace couleur par l'analyse de la représentation de la couleur de peau dans les différents espaces. Nous avons choisi l'espace YCbCr qui permet une représentation stable et compacte par une distribution gaussienne d'un modèle couleur. Nous avons ensuite étudié la meilleure méthode de définition du modèle couleur. Nous avons choisi un modèle local à partir d'une fenêtre échantillon plutôt qu'un modèle global par un apprentissage à partir d'une base de référence. Cette approche permet une définition adaptée à l'image à traiter et permet de prendre en compte la grande variété de teinte possible de la chevelure. Enfin nous avons défini notre classifieur couleur en comparant plusieurs algorithmes de segmentation de la couleur de peau. Nous avons choisi un modèle couleur par distribution gaussienne 1D pour chacune des composantes de chrominance de l'espace YCbCr.



FIGURE 5.12 – Analyse Couleur - Résultats : Exemples de cartes couleurs

6 Fusion de données

Sommaire

6.1	La Fusion de Données	121
6.1.1	Processus	121
6.1.2	Les Sources	122
6.1.3	Les Approches	123
6.1.3.1	Méthodes Classiques	123
6.1.3.2	Méthodes liées à un traitement non probabiliste de l'incertain	124
6.2	Les Fonctions de croyances	126
6.2.1	Le Modèle de <i>Croyances Transférables</i>	126
6.2.1.1	Formalisme du MCT	126
6.2.1.2	La fonction de Masse	127
6.2.1.3	La Fonction de répartition de Masse (FM)	127
6.2.1.4	Transformation équivalente de la MCT	128
6.2.2	Règles de Combinaison	128
6.2.2.1	Combinaison Conjonctive	129
6.2.2.2	Combinaison Disjonctive	129
6.2.2.3	Propriétés de combinaisons	129
6.2.3	Prise de Décision	130
6.2.3.1	Prise de décision à partir des probabilités pignistiques	130
6.2.3.2	Utilisation des plausibilités pour la prise de décision	130
6.2.3.3	Modélisation de l'incertitude	131
6.2.4	Gestion du Conflit	131
6.3	La Fusion Capillaire	132
6.3.1	Fusion Logique	132
6.3.2	Le MCT Capillaire	134
6.3.2.1	Les 3 espaces de fusion de croyances	135
6.3.2.2	Source Fréquentielle et Fonction de masse m_f	136
6.3.2.3	Source Couleur et Fonction de masse m_c	137
6.3.2.4	DRC Capillaire	138
6.3.2.5	CRC Capillaire	140
6.3.2.6	Prise de Décision Capillaire	140
6.3.2.7	Résultats fonctions de croyances	142
6.3.3	Fonction d'affaiblissement	142
6.3.3.1	Formalisme	142
6.3.3.2	Résultats	145
6.3.3.3	Pondération de la fusion Logique	149
6.4	Conclusion	150

Nous avons maintenant défini les masques de localisation fréquentielle et colorimétrique qui représentent les zones les plus probables de la position de la chevelure. Il nous faut maintenant réunir l'information commune afin d'initialiser au mieux le placement des marqueurs pour le traitement "*Matting*". Une fois les marqueurs placés par notre processus de fusion de données, nous obtenons un masque capillaire après estimation de l'alpha matte par le traitement "*Matting*". Au final, nous comparons les masques capillaires segmentés aux masques de références pour évaluer les résultats obtenus.

Dans ce chapitre, nous présenterons dans un premier temps, le procédé de fusion de données (*section 6.1*). Dans un second temps, nous introduirons en détail la théorie des fonctions de croyances (*section 6.2*). Puis nous présenterons l'approche mise en place pour la segmentation capillaire (*section 6.2*). Nous commencerons par appliquer une fusion logique des données fréquentielles et colorimétriques (*section 6.3.1*) puis nous appliquerons un modèle de fonction de croyances adapté à notre segmentation (*section 6.3.2*). Pour finir nous optimiserons notre algorithme avec la prise en compte d'information dite de "haut niveaux" comme la localisation de la chevelure par rapport au visage par la mise en place d'une fonction d'affaiblissement (*section 6.3.3*)

6.1 La Fusion de Données

La mesure de grandeur physique par un capteur est soumise à la fiabilité du capteur ainsi qu'aux conditions d'acquisition, ce qui entraîne une certaine incertitude.

« La fusion de données vise à l'association, la combinaison, l'intégration et le mélange de multiples sources de données représentant des connaissances et des informations diverses dans le but de fournir une meilleure décision par rapport à l'utilisation séparée des sources de données ».

6.1.1 Processus

Dans un processus de fusion, quatre phases principales sont enchaînées successivement :

- **Représentation homogène et recalage des informations pertinentes** : les données à fusionner sont souvent hétérogènes, il est impossible de les combiner sous leur forme initiale. On est alors amené à rechercher un espace de représentation commun dans lequel les différentes informations pertinentes disponibles renseignent sur une même entité. Un premier traitement consiste donc à transformer certaines de ces informations initiales, en informations équivalentes dans un espace commun, dans lequel s'effectuera la fusion.
- **Modélisation des connaissances** : chaque jeu de données propre à chaque source n'est pas forcément exploitable en tant que tel, notamment si l'information fournie est très imparfaite et ne donne qu'un aspect de la réalité. Cependant, même imparfaite, toute information peut apporter de la connaissance sur l'état du système. Donc, une étape essentielle du processus de fusion consiste à modéliser et à évaluer la connaissance apportée par chaque source. Elle est couplée au choix d'un cadre théorique adapté.
- **Fusion** : c'est à ce niveau du processus que l'opération de fusion proprement dite est réalisée. Les informations recalées et modélisées sont combinées selon une règle de combi-

raison propre au cadre théorique choisi. Observons notamment que la fusion d'informations contradictoires doit permettre de gérer les conflits potentiels entre les diverses sources.

- **Décision par choix d'une stratégie** : la fusion doit permettre de choisir l'information la plus vraisemblable, au sens d'un certain critère, parmi toutes les hypothèses possibles. En ce sens, la fusion de données aboutit bien souvent à une classification (affectation d'un ensemble de mesures aux hypothèses possibles). Le critère de décision dépend du cadre théorique dans lequel le processus de fusion a été développé et de l'objectif à atteindre. En particulier, selon le niveau de précision et/ou de certitude recherché.

Trois éléments sont toujours considérés dans le cadre de la fusion de données : les sources, les algorithmes et le résultat. Le résultat sera composé de données ayant une valeur ajoutée par rapport aux données initiales en provenance de la source. La nature de cette valeur ajoutée dépendra de l'application.

6.1.2 Les Sources

La principale difficulté de la fusion de données est de pouvoir modéliser et évaluer la connaissance apportée par chaque source afin de tirer le meilleur parti des informations disponibles. Cette opération consiste à exprimer l'imperfection de chaque information en fonction de la source dont elle est issue.

- **Les niveaux de fusion** :

D'après Hall *et al* [Hall 01], on peut définir trois types de bases de fusion :

- La fusion de mesures qui combine directement les mesures de chaque capteur, elle est encore appelée fusion *amont*.
- La fusion de primitives qui combine des caractéristiques extraites par chaque capteur.
- La fusion de décisions qui combine l'information après que chaque capteur ait déterminé une localisation, une identité partielle des objets présents, elle est encore appelée fusion *aval*.

- **La relation entre les sources** :

Connaître les relations entre les sources est important car ceci permettra de caractériser, d'ordonner ou encore de corrélérer les différentes sources de données disponibles. On pourra alors décider quelle est la meilleure stratégie ou le meilleur algorithme pour fusionner les données. En se basant sur les travaux de Bellot [Bellot 02], nous présentons les relations suivantes entre les sources :

- **distribution** : les sources de données donnent des informations sur le même environnement mais chacune en ayant soit un point de vue différent, soit la capacité de n'observer qu'une partie de l'environnement.
- **complémentarité** : chaque source perçoit uniquement un sous-ensemble de l'environnement global. La fusion de l'ensemble de ces sources donnera une vue élargie de l'environnement. Les sources, dans ce cas, ne sont pas redondantes.
- **hétérogénéité** : chaque source fournit des informations dont les caractéristiques sont complètement différentes des autres sources.
- **redondance** : les sources de données décrivent le même environnement avec des informations de même caractéristique. C'est le cas lorsqu'un capteur est doublé afin de

prévenir une panne ou simplement pour augmenter la précision ou la confiance accordée aux mesures.

- **contradiction** : l'espace de définition de deux sources est le même mais elles fournissent des informations totalement différentes l'une de l'autre, tout en observant le même environnement.
- **concordance** : les données fournies par deux sources sont compatibles entre elles et se corroborent l'une l'autre.
- **discordance** : les informations fournies par deux sources sont incompatibles entre elles. La différence entre la contradiction et la discordance est qu'ici une source ne va pas invalider l'information donnée par l'autre source, mais simplement donnera un résultat qui est aussi acceptable (dans le cadre de l'application) que celui de la première source.
- **différence de granularité** : les sources de données fournissent des données redondantes mais chaque source observe l'environnement à une échelle différente.
- **synchronisation** : les données fournies par les sources sont concordantes au niveau du temps.

Cette liste présente donc les caractéristiques des relations pouvant exister entre des sources de données. Cependant, il apparaît clairement que dans une application réelle, les sources en présence pourront être munies de plusieurs de ces relations.

6.1.3 Les Approches

La fusion est un procédé de combinaison de données et d'information pour estimer ou prédire des états. La combinaison des données vise essentiellement à réduire l'incertitude sur l'information. Les incertitudes peuvent être modélisées par différentes théories, nous présenterons une liste non exhaustive des principales approches de fusion.

6.1.3.1 Méthodes Classiques

La mesure de probabilité : L'utilisation de raisonnement de type Bayésien basé sur la théorie des probabilités fait partie des méthodes de prise de décision les plus anciennes. Dans cette approche, les capteurs sont considérés comme un ensemble d'entités capables de fournir une décision à tout instant. Chaque capteur est alors vu comme un estimateur bayésien.

Ainsi, les distributions de probabilités associées à chaque capteur sont combinées dans une seule fonction de distribution de probabilités jointes a posteriori, en utilisant la règle de Bayes pour l'ensemble H des possibilités exhaustives $H = \{h_1 \dots h_n\}$ pour l'ensemble O des observations $O = \{o_1 \dots o_n\}$

$$P(h_i|o_j) = \frac{P(h_i) \cdot P(o_j|h_i)}{\sum_{i=1}^n P(h_i) \cdot P(o_j|h_i)}, \forall h_i \in H \quad (6.1)$$

La vraisemblance de cette fonction est maximisée pour obtenir la fusion finale de l'ensemble des capteurs. Les systèmes probabilistes sont à la base de nombreux systèmes actuels, par exemple pour le traitement de signaux audio et vidéo. Cette approche permet une modélisation relativement simple, mais elle ne permet pas de représenter des modèles complexes évoluant dans le temps.

Utilisation d'Estimateurs : Dans ce contexte, les données récupérées à partir des capteurs et les erreurs de mesures associées sont interprétées comme un ensemble de contraintes sur l'espace des solutions. Chaque point de l'espace représente une position possible pour l'objet observé. Au fur et à mesure que les données arrivent, l'espace des solutions est réduit itérativement, jusqu'à un point de convergence qui représente alors la position réelle de l'objet.

L'utilisation d'estimateurs s'applique aux données de bas niveau : ils permettent de convertir des données observées en estimations de l'état d'une cible par exemple. Il s'agit donc plus spécifiquement de méthodes de reconstruction de l'environnement. Les problèmes liés à la localisation statique sont résolus par des estimateurs qui, à partir de données multiples prises aux mêmes instants, fournissent une estimation de position, par contre, dans les cas de suivi dynamique, ces estimateurs s'appliquent à des séquences temporelles de données. L'estimateur statistique le plus utilisé est le filtre de Kalman qui fournit une solution récursive aux problèmes d'estimation à variance minimale.

6.1.3.2 Méthodes liées à un traitement non probabiliste de l'incertain

Les probabilités sont l'une des mesures de confiance possibles. On peut concevoir d'utiliser pour la fusion d'autres mesures non additives, qui peuvent donner lieu à des combinaisons et des agrégations de types plus variés et supposent moins de connaissance a priori sur les données.

La Théorie des "Possibilités" : la théorie des possibilités est une extension de la théorie des ensembles flous introduite par Zadeh [Zadeh 65] et développée par Dubois *et al* [Dubois 88]. Ces travaux montrent que les mesures de possibilités permettent de représenter facilement des notions statistiques complexes. Dans la théorie classique des ensembles, un élément appartient à l'ensemble A (un ensemble de nombre réels), si sa fonction d'appartenance μ vaut 1, et 0 dans le cas contraire. Cette notion est étendue au cas continu en définissant une fonction d'appartenance continue dans $[0,1]$. En d'autres termes, un élément peut appartenir à A avec un degré d'appartenance non Booléen. On définit alors la distribution de possibilité π associée à A telle que π corresponde à la fonction d'appartenance de A .

Formellement, une telle distribution fournit une double mesure de la vraisemblance d'un événement, à savoir la mesure de "*Possibilité*" Π et celle de "*Nécessité*" N .

La mesure de "*Possibilité*" Π est définie par l'équation :

$$\forall A, B \in \Omega \quad \Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \quad (6.2)$$

La mesure de "*Nécessité*" N est elle définie par l'équation :

$$\forall A, B \in \Omega \quad N(A \cap B) = \min(N(A), N(B)) \quad (6.3)$$

Ces mesures sont les seules, à pouvoir être représentée par une distribution, c'est-à-dire par des valeurs sur les singletons de Ω . Cette distribution est notée π , elle prend ses arguments dans Ω et est à ses valeurs dans $[0,1]$. La "*Possibilité*" et la "*Nécessité*" sont alors définies à partir de la distribution π par les équations :

$$\begin{aligned} \Pi(A) &= \sup_{x \in A} \pi(x) \quad \forall A \in \Omega \\ N(A) &= 1 - \Pi(A^c) = \inf_{x \in A} \pi(x) \quad \forall A \in \Omega \end{aligned} \quad (6.4)$$

Soient deux distributions de possibilité π_1 et π_2 (figure 6.1a). On peut fusionner de trois manières différentes ces distributions de "Possibilité" :

- **Fusion conjonctive** : ce premier type comprend l'intersection des distributions de "Possibilité" des sources qui donne donc une idée pessimiste de la combinaison (T-normes) (figure 6.1b)
- **Fusion disjonctive** : Ce premier type comprend l'union de ces distributions qui donne une vision trop optimiste (T-conormes) (figure 6.1c)

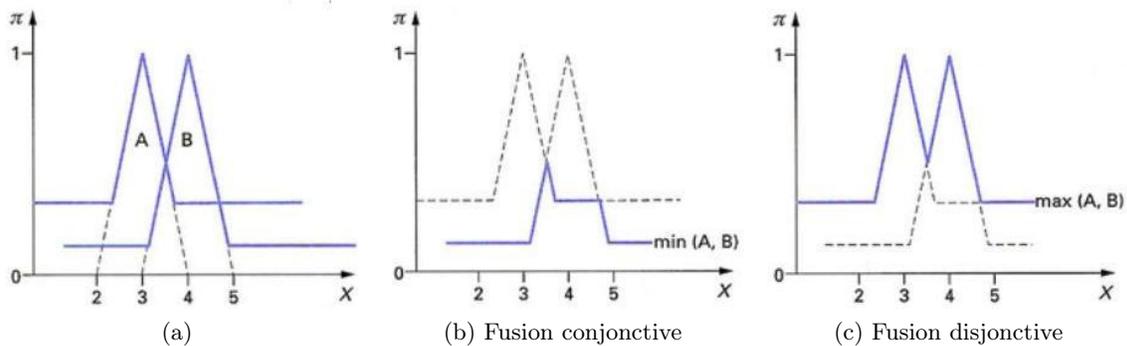


FIGURE 6.1 – La Théorie des "Possibilités" : Schéma des fusions conjonctive et disjonctive

- **Fusion adaptée** : elle permet de gérer le conflit entre les distributions, lorsque les noyaux des distributions ne se chevauchent pas. La mesure du conflit est liée au degré de consensus h entre π_1 et π_2 (figure 6.2a) et la borne supérieure de leur intersection est égale à $1 - h$. La règle de combinaison a été proposée par Dubois *et al* [Dubois 92] (figure 6.2b).

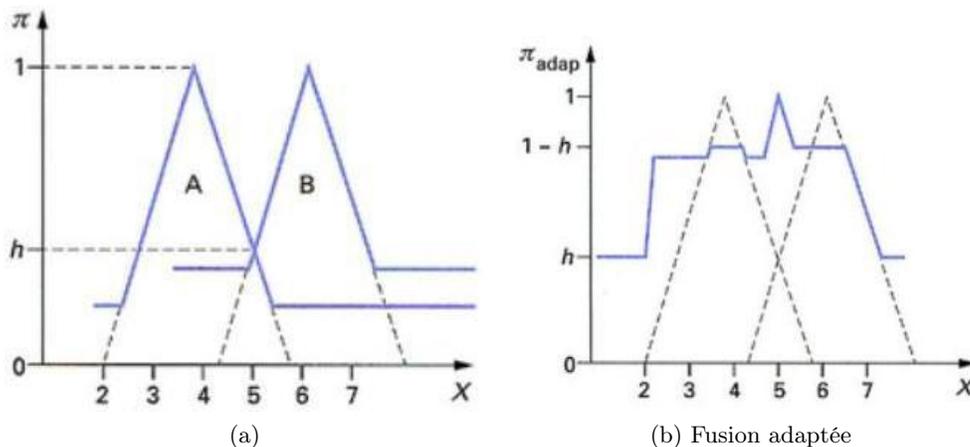


FIGURE 6.2 – La Théorie des "Possibilités" : Schéma de la fusion adaptée

Les opérateurs dépendant du contexte sont quant à eux beaucoup plus puissants car, selon le conflit existant entre les sources ou la fiabilité de celles-ci, on module leur comportement : par exemple, si les sources présentent un consensus, on réalisera une conjonction, s'il y a conflit, on réalisera une disjonction et si le conflit est faible, on réalisera un compromis.

La Théorie des Fonctions de Croyances ou de l'évidence : Cette théorie permet de représenter l'ignorance de façon explicite en se dispensant de la clause d'équirépartition qui existe

avec une mesure additive.

Cette théorie est dérivée de l'approche bayésienne, mais utilise deux mesures pour qualifier le degré de croyance que l'on a sur une hypothèse, calculé à partir d'indices la confirmant ou l'infirmité. La théorie peut assigner une mesure de certitude à des ensembles d'hypothèses autant qu'à des hypothèses seules. Cette approche permet de raisonner sur des ensembles d'hypothèses dans un premier temps et de se restreindre petit à petit aux hypothèses plausibles, au fur et à mesure que de nouvelles évidences apparaissent. Cette approche de la fusion de données est adaptée à la fusion de multiples capteurs. Dans de nombreuses applications, cette modélisation de l'ignorance est en effet nécessaire pour déterminer la vraisemblance d'hypothèses obtenues par des capteurs délivrant des informations incomplètes ou à un niveau sémantique plus élevé.

6.2 Les Fonctions de croyances

La théorie Dempster [Dempster 77], complétée par les propositions Shafer [Shafer 76], a donné naissance à un modèle mathématique de croyance connue sur le nom de la théorie de l'évidence. Elle permet de représenter explicitement, à partir d'outils mathématiques, l'incertitude liée aux connaissances (tels que l'avis des experts sur le degré de plausibilité ou de croyance que la pluie tombera dans 5 jours). Par cette méthode, on peut modéliser le degré de "croyance" qui peut se révéler particulièrement efficace lors de la fusion de différents points de vue.

Connue pour son efficacité dans de multiples applications tels que : la classification d'images, prise de décision, nous présentons la théorie de Dempster-Shafer à travers ses principaux termes fondamentaux, afin de permettre d'en saisir le sens. Cependant, il existe plusieurs interprétations des fonctions de croyances : le modèle de Dempster [Dempster 77], Le modèle de Shaffer [Shafer 76] ou le modèle des Croyances Transférables [Smets 94] par exemple. Nous nous plaçons dans ce dernier pour la présentation du formalisme mathématique dans le cadre axiomatisé et formalisé par Smets *et al* [Smets 94] [Rammasso 07].

6.2.1 Le Modèle de *Croyances Transférables*

Le *Modèle de Croyances Transférables* MCT (TBM : *Transferable Belief Model*) se divise en 2 grandes étapes :

- Le niveau *crédal* (du latin "credo" se traduisant "je crois") qui représente le niveau de la représentation et de la combinaison de l'information
- Le niveau *pignistique* (du latin "pignus" se traduisant "un pari") qui représente le niveau de prise de décision

Cette représentation en 2 étapes se différencie des autres approches, en particulier celle de la théorie des probabilités ou de la représentation des connaissances, par la combinaison et la décision effectuées sans distinguer les niveaux d'abstraction. Contrairement à la théorie des probabilités, elle permet en particulier de modéliser différents types d'ignorance.

6.2.1.1 Formalisme du MCT

L'état du système étudié est supposé prendre des valeurs discrètes appelées hypothèses ω_k et l'ensemble des K hypothèses possibles décrivant les différents états du système est appelé *Cadre* ou *Univers* que l'on décrit selon l'équation :

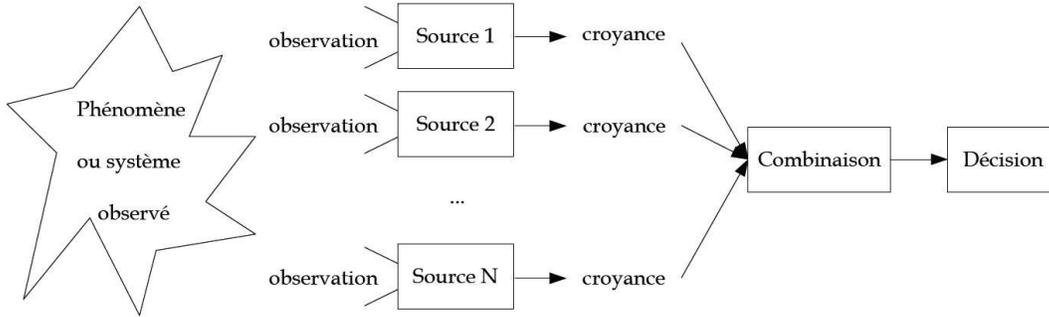


FIGURE 6.3 – **Fusion d'information** : des sources de croyances émettent une opinion pondérée sur l'état du système observé. Les différentes sources sont combinées afin de synthétiser la compréhension du système et améliorer la décision sur l'état du système après combinaison [Rammasso 07].

$$\Omega = \{\omega_1 \dots \omega_k \dots \omega_K\} = \bigcup_{k=1}^K \{\omega_k\} \quad (6.5)$$

Le nombre d'hypothèses de $|\Omega| = K$ composant Ω est appelé cardinal et les hypothèses sont considérées exclusives.

L'état réel du système $\omega_r \subseteq \Omega$ est décrit par les différentes sources d'information décrivant l'état du système. L'information apportée n'est pas fiable à 100%. On définit ainsi une croyance sur l'état du système.

6.2.1.2 La fonction de Masse

La croyance peut être représentée sous plusieurs formes. La représentation la plus couramment utilisée est la fonction de masse de croyances définie pour une source de croyances S :

$$m_S^\Omega : \begin{array}{l} 2^\Omega \rightarrow [0, 1] \\ B \mapsto m_S^\Omega(B) \end{array} \quad (6.6)$$

où l'espace puissance $2^\Omega = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_3\}, \dots, \{\omega_1, \dots, \omega_K\}\}$ décrit l'ensemble des sous-ensembles des états possibles de Ω et B une proposition de Ω telle que $B \in 2^\Omega$.

6.2.1.3 La Fonction de répartition de Masse (FM)

Une distribution de masses de croyances est définie pour une source S est l'ensemble des masses de croyances tel que :

$$\sum_{B \subseteq \Omega} m_S^\Omega(B) = 1 \quad (6.7)$$

La masse $m_S^\Omega(B)$ représente le part de croyance de la source S en la proposition " $\omega_r \in B$ " où ω_r est l'état réel du système observé. Un élément B dont la masse est non nulle (ie $m_S^\Omega(B) > 0$) est appelé élément focal (EF) de la FM.

Une fonction de masse peut être qualifiée de :

Fonction		Rôles
Fonction de masses	m	Représentation de la connaissance
Plausibilité	pl	Vraisemblances
Crédibilité	bel	Représentation de la connaissance
Implicabilité	b	Règle de Combinaison Disjonctive (DRC)
Communalité	q	Règle de Combinaison Conjonctive (CRC)

TABLE 6.1 – "Fonctions de Croyances"

- **FM normale** si l'ensemble vide \emptyset n'est pas un EF ie $m^\Omega(\emptyset) = 0$
- **FM catégorique** si elle possède un unique EF
- **FM vide** si Ω est l'unique élément focale ie $m^\Omega(\Omega) = 1$. Cela traduit l'ignorance totale de la source sur l'état ω_r du système
- **FM dogmatique** si Ω n'est pas un EF ie $m^\Omega(\Omega) = 0$

6.2.1.4 Transformation équivalente de la MCT

À partir de la représentation de la fonction de masse (FM) m^Ω , d'autres représentations de la même information sont possibles. Elles sont toutes définies de 2^Ω vers $[0,1]$ cependant la somme des EF n'est pas forcément unitaire.

- **Plausibilité** pl : la plausibilité mesure la confiance maximum que l'on peut avoir en B.

$$pl^\Omega(B) = \sum_{C \cap B \neq \emptyset} m^\Omega(C), \forall B \subseteq \Omega \quad (6.8)$$

- **Crédibilité** bel : la crédibilité mesure la force avec laquelle on croit à la proposition B

$$bel^\Omega(B) = \sum_{\emptyset \neq C \subseteq B} m^\Omega(C), \forall B \subseteq \Omega \quad (6.9)$$

- **Implicabilité** b : l'implicabilité représente la somme des masses allouées aux sous-ensembles de B tels que leur véracité implique la véracité de B

$$b^\Omega(B) = \sum_{C \subseteq B} m^\Omega(C) = bel^\Omega(B) + m^\Omega(\emptyset), \forall B \subseteq \Omega \quad (6.10)$$

- **Communalité** q : la communalité représente la somme des masses allouées aux sur ensemble de B et donc qui ont en commun B

$$q^\Omega(B) = \sum_{C \supseteq B \neq \emptyset} m^\Omega(C), \forall B \subseteq \Omega \quad (6.11)$$

6.2.2 Règles de Combinaison

Nous venons de décrire comment définir une fonction de croyances pour une source d'information, il nous faut maintenant définir les règles de combinaison qui gèrent la fusion de données pour déterminer la prise décision. La théorie de Dempster-Shafer permet de combiner des informations issues de sources différentes, car d'après Shafer, si deux fonctions de croyances sont définies sur un même cadre de discernement, alors ce cadre de discernement distingue l'interaction pertinente entre ces deux informations. De ce fait les sources doivent être définies sur un

même cadre de discernement de façon que l'on puisse en déduire leur somme orthogonale suivant les règles de combinaisons de Dempster. Cette somme est toujours une fonction de croyances et prend en compte l'influence de toutes les autres.

Nous présenterons les 2 règles de combinaisons basées sur l'hypothèse de sources de croyances distinctes.

6.2.2.1 Combinaison Conjonctive

Soient 2 FM m_1^Ω et m_2^Ω distinctes définies pour le même univers Ω dont les sources 1 et 2 sont supposées fiables [Shafer 76]. La règle de Combinaison Conjonctive (CRC : Conjonctive Rule of Combination), noté \odot , des FM m_1^Ω et m_2^Ω ou de leur communalité associée q_1^Ω et q_2^Ω sont définies $\forall B, C, D \subseteq \Omega$ par les équations suivantes :

$$m_{1\odot 2}^\Omega(D) = (m_1^\Omega \odot m_2^\Omega)(D) = \sum_{B \cap C = D} m_1^\Omega(B) \cdot m_2^\Omega(C) \quad (6.12)$$

$$q_{1\odot 2}^\Omega(D) = (q_1^\Omega \odot q_2^\Omega)(D) = q_1^\Omega(D) \cdot q_2^\Omega(D) \quad (6.13)$$

L'intersection intervenant dans la CRC joue un rôle de *spécialisation* en transférant la masse sur des sous-ensembles de cardinalité plus forte.

6.2.2.2 Combinaison Disjonctive

Soient 2 FM m_1^Ω et m_2^Ω distinctes définies pour le même univers Ω dont au moins l'une de ces sources 1 ou 2 est supposée fiable sans pour autant connaître laquelle ni en quantifier la fiabilité. [Shafer 76]. La règle de Combinaison Disjonctive (DRC : Disjonctive Rule of Combination), noté \oplus , des FM m_1^Ω et m_2^Ω ou de leur implicabilité associée b_1^Ω et b_2^Ω est définie $\forall B, C, D \subseteq \Omega$ par les équations suivantes :

$$m_{1\oplus 2}^\Omega(D) = (m_1^\Omega \oplus m_2^\Omega)(D) = \sum_{B \cup C = D} m_1^\Omega(B) \cdot m_2^\Omega(C) \quad (6.14)$$

$$b_{1\oplus 2}^\Omega(D) = (b_1^\Omega \oplus b_2^\Omega)(D) = b_1^\Omega(D) \cdot b_2^\Omega(D) \quad (6.15)$$

L'intersection intervenant dans la DRC joue un rôle de *généralisation* en transférant la masse sur les sous ensembles de cardinalité plus faible.

6.2.2.3 Propriétés de combinaisons

Les propriétés notables des règles de combinaisons sont les suivantes :

- La CRC et la DCR sont des règles associatives commutatives et non idempotentes (ie $(m_1^\Omega \odot m_1^\Omega) \neq m_1^\Omega$) et par conséquent il est impossible de combiner des sources dépendantes.
- L'élément neutre est la FM vide et $m^\Omega(\emptyset) = 1$ est l'élément absorbant.

- La dualité des règles \odot et \oplus est mise en valeur par les règles de De Morgan s’appliquant aux fonctions de croyances.
- La CRC et la DCR sont des opérations linéaires qui peuvent être mises sous forme matricielle.
- L’utilisation des communalités (respectivement implicabilité) réduit la CRC (DCR) à seulement un produit membres à membres

6.2.3 Prise de Décision

Les systèmes de fusion d’information, qu’ils soient basés sur la théorie des probabilités, des possibilités ou des fonctions de croyances, ont pour finalité la prise de décision. Prendre une décision consiste à choisir une hypothèse sur un cadre de pari, généralement le cadre de discernement. La prise de décision peut être réalisée de façon automatique ou laissée à la responsabilité de l’utilisateur final (par exemple dans le cas de l’aide au diagnostic dans le domaine de la Médecine).

6.2.3.1 Prise de décision à partir des probabilités pignistiques

Dans le cadre du TBM, la phase de décision s’appuie sur la distribution de probabilités pignistiques [Smets 05] notée $BetP\{m\}$ obtenue à partir de la FM m . La transformée pignistique consiste à répartir de manière équiprobable la masse d’une proposition B sur les hypothèses contenues dans B . Formellement :

$$BetP\{m^\Omega\} : \Omega \rightarrow [0, 1] \quad (6.16)$$

$$\omega_k \mapsto BetP\{m^\Omega\}(\omega_k) \quad (6.17)$$

où $BetP\{m^\Omega\}(\omega_k)$, $\forall \omega_k \in \Omega$ est donnée par :

$$BetP\{m^\Omega\}(\omega_k) = \frac{1}{(1 - m^\Omega(\emptyset))} \sum_{B \subseteq \Omega, \omega_k \in B} \frac{m^\Omega(B)}{|B|} \quad (6.18)$$

La décision est généralement prise en choisissant l’élément ω_k possédant la plus grande probabilité pignistique :

$$\omega_r = \arg \max_{\omega_k \in \Omega} BetP\{m^\Omega\}(\omega_k) \quad (6.19)$$

6.2.3.2 Utilisation des plausibilités pour la prise de décision

L’agent qui se base sur la transformée pignistique lors de la phase de prise de décision a un comportement rationnel en maximisant l’utilité espérée. Alternativement, les plausibilités et les crédibilités peuvent être utilisées lorsque l’agent a une attitude plutôt optimiste ou pessimiste. Dans les cas de la plausibilité, on dit que l’agent a une attitude optimiste dans le sens où la plausibilité d’un élément représente la valeur maximale de la masse de croyances pour cet élément qui pourrait être atteinte si des informations supplémentaires parvenaient au système de fusion. La distribution de probabilités sur les singletons à partir de laquelle une décision peut être prise est alors donnée $\forall \omega_k \in \Omega$ par l’équation :

$$PlP\{m^\Omega\}(\omega_k) = \frac{pl^\Omega(\omega_k)}{\sum_{\omega_k \in \Omega} pl^\Omega(\omega_k)} \quad (6.20)$$

Notons que la plausibilité d'un singleton $pl^\Omega(\omega_k)$ est égale à sa communalité $q^\Omega(\omega_k)$

6.2.3.3 Modélisation de l'incertitude

À partir de la distribution des masses de croyances, les bornes supérieures et inférieures d'un intervalle de probabilité peuvent être définies. Cet intervalle contient la probabilité d'un ensemble d'hypothèses ou éléments focaux et est borné par les deux mesures non-additives de crédibilité bel et plausibilité pl .

Il en résulte la propriété d'encadrement définie pour tout élément $B \in 2^\Omega$ selon l'équation suivante :

$$bel(B) \leq BetP\{m^\Omega\}(B) \leq pl(B) \quad (6.21)$$

Où $BetP\{m^\Omega\}(B)$ exprime la probabilité que B se réalise mais qui reste inconnue et peut prendre toute valeur dans $[bel(B), pl(B)]$.

Les mesures de plausibilité et de crédibilité ne sont pas duales car elles ne sont pas additives au sens de la théorie des probabilités $bel(B) \neq pl(B^c)$ ou B^c est le complément de B par rapport à Ω . Toutefois, les relations ci-dessous peuvent être établies entre B et B^c :

$$\begin{aligned} bel(B) &= 1 - pl(B^c) \\ pl(B) &= 1 - bel(B^c) \end{aligned} \quad (6.22)$$

L'incertitude est décrite par la différence $(bel(B) - pl(B))$ figure(6.4) concernant l'hypothèse B représentée par l'intervalle $[bel(B), pl(B)]$

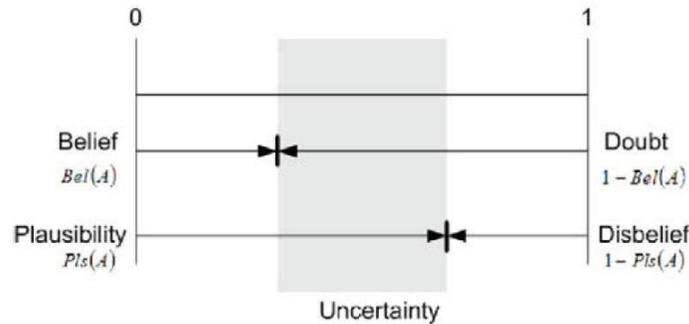


FIGURE 6.4 – Modélisation de L'incertitude

6.2.4 Gestion du Conflit

Lors de la combinaison par la CRC de deux FM m_1^Ω et m_2^Ω , l'intersection entre les éléments focaux peut être vide et dans ce cas la FM résultante a une masse non nulle sur l'ensemble vide, i.e. $m_1^\Omega \circledast m_2^\Omega(\emptyset) > 0$. Cette valeur est appelée conflit et quantifie la discordance entre les sources de croyances. De plus l'ensemble vide étant absorbant par l'opérateur CRC, la combinaison avec une troisième FM m_3 ne fera qu'augmenter la valeur du conflit $m_1 \circledast m_2 \circledast m_3(\emptyset) > m_1 \circledast m_2(\emptyset)$.

Il est nécessaire d'analyser les origines du conflit qui peuvent être multiples [Smets 05] :

- **La non exhaustivité du cadre de discernement** : la quantité de masse de conflit peut être interprétée comme la croyance sur une nouvelle hypothèse : on parle alors de TBM

en monde ouvert.

- **La nature de la combinaison conjonctive** : la CRC n'est pas idempotente et par conséquent la combinaison d'une FM avec elle-même modifie la FM initiale et peut induire du conflit qui est alors appelé conflit interne. Pour éviter ce phénomène, deux solutions existent. La première consiste à construire des distributions de masses consonantes en s'inspirant des sous-ensembles flous. La seconde consiste à changer la règle de combinaison conjonctive et à utiliser une règle minimisant le conflit.
- **Une nouvelle source** : le conflit peut être aussi interprété comme une véritable source d'information par exemple pour regrouper des fonctions de croyances.

Dans certaines applications, il peut être intéressant de traiter le conflit automatiquement et pour cela il existe des règles dites de redistribution. Celle de Yager par exemple transfère le conflit sur le doute total (Ω) [Yager 87] tandis que la règle de Dubois *et al* [Dubois 88], aussi appelée règle adaptative, transfère le conflit sur l'union des propositions impliquées.

6.3 La Fusion Capillaire

Notre processus de fusion consiste à utiliser l'information acquise par nos analyses fréquentielles et couleur afin de conserver l'information utile pour la localisation des marqueurs du traitement "*Matting*". Nous cherchons à positionner les marqueurs cheveux H_m et les marqueurs non cheveux \bar{H}_m . Les zones non marquées correspondent aux zones d'incertitude où l'état des pixels sera donné par l'estimation de l'alpha matte.

6.3.1 Fusion Logique

Nous présentons dans cette section une approche simple par fusion logique à partir des cartes binaires de l'analyse fréquentielle et colorimétrique. Nous combinons l'information des deux cartes selon les règles de fusion définies dans le tableau 6.2 :

Carte binaire fréquentielle	Carte binaire couleur	fusion logique	marqueurs " <i>Matting</i> "
0	0	0	\bar{H}_m
0	1	0	\bar{H}_m
1	0	\emptyset	Un
1	1	1	H_m

TABLE 6.2 – Table des règles du modèle de fusion logique

Les marqueurs cheveux H_m sont placés pour les pixels appartenant au masque de fusion logique ($F = 1$ & $C = 1$). Les marqueurs non cheveux \bar{H}_m sont placés à l'extérieur du masque binaire fréquentielle ($F = 0$) pour les pixels dans ($C = 1$) et hors ($C = 0$) de la carte couleur. L'information portée par le masque binaire de couleur n'est pas suffisamment pertinente pour admettre que les pixels dans le masque fréquence ($F = 1$) et hors du masque de couleur $C = 0$) soient des pixels appartenant à la zone cheveux ou non cheveux. Ces pixels sont considérés comme des pixels dont l'état est "inconnu" (similaire à l'état incertain Un) et seront estimés pendant le traitement "*Matting*".

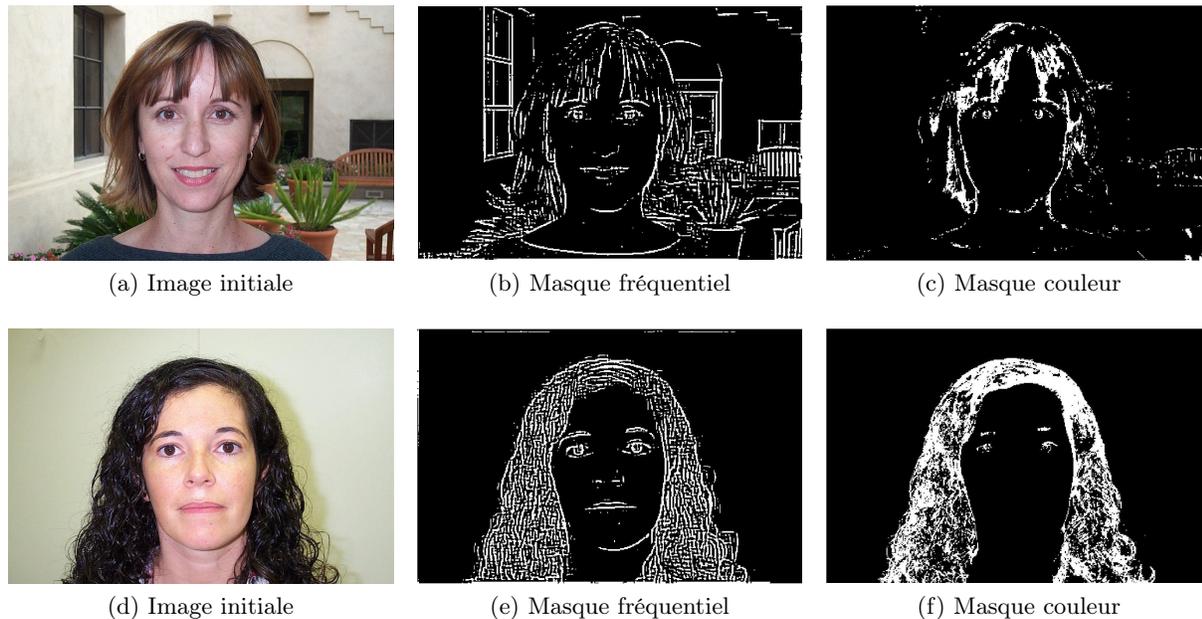


FIGURE 6.5 – **Fusion Logique** : Exemples de masques binaires de localisation des cheveux en fréquence et en couleur pour la fusion logique de données. On représente les pixels blancs pour le masque binaire= 1 et les pixels noirs pour le masque binaire= 0.

Nous évaluons notre approche par fusion logique pour 200 images de la base maquillage et 160 images de la base Caltech. Les résultats sont calculés sur les masques capillaires binaires obtenus après segmentation par le traitement "*Matting*" initialisé par les marqueurs définis par fusion logique.

	T.F.R.	$\sigma(\text{T.F.R.})$	T.F.A.	$\sigma(\text{T.F.A.})$
Fusion Logique (base maquillage)	9.1%	12.7%	14.2%	15.4%
Fusion Logique (base Caltech)	6.7%	16.3%	40.6%	21.6%

TABLE 6.3 – Taux de faux rejet (T.F.R.) et de fausse acceptance (T.F.A.) pour la segmentation capillaire par fusion logique. Résultats obtenus pour 200 images de la base maquillage et 160 images de la base Caltech

Cette première approche va nous permettre d'évaluer la pertinence de nos études fréquentielle et couleur. En effet les marqueurs cheveux ne dépendent que du masque de fusion entre les deux cartes binaires de localisation. Une bonne segmentation permettra de valider notre choix d'utiliser l'information fréquentiel et couleur pour la localisation d'une chevelure dans une image.

Pour les premiers résultats obtenus par la fusion logique, nous obtenons un faible taux de faux rejet de seulement 9.1% pour la base maquillage et de 6.7% pour la base Caltech. Pour ces deux bases, le taux de faux rejet est faible ce qui implique que la chevelure est généralement bien segmentée avec très peu de perte des pixels cheveux avec un faible écart type moyen pour les 2 bases. Ces résultats s'expliquent généralement pour plusieurs raisons :

- *la sous-segmentation des pixels cheveux à la frontière du visage* : Les pixels des racines des

cheveux en périphérie du visage sont peu marqués H_m et la proximité des marqueurs non cheveux \bar{H}_m des pixels de peau entraînent la mauvaise segmentation de ces pixels.

- *la sous-segmentation des cheveux longs* : Comme pour les racines des cheveux, les cheveux long peuvent être insuffisamment marqués. Les marqueurs cheveux H_m sont localisés trop loin des pointes où il peut y avoir des marqueurs non cheveux \bar{H}_m .
- *la sous-segmentation de certaines mèches* : La mauvaise localisation des marqueurs cheveux H_m et non cheveux \bar{H}_m peut entraîner le non marquage de certaines zones. En particulier les mèches de cheveux avec beaucoup de reflet spéculaire où l'information couleur et fréquentielle est saturée.

Pour le taux de fausse acceptance, nous observons un faible taux de 14.2% pour la base maquillage et plus important de 40.6% pour la base Caltech. La sur-segmentation de la chevelure est assez faible pour la base maquillage car les modèles sont pris sur des fonds uniformes, ainsi seuls les vêtements peuvent être sur-segmentés. Pour la base Caltech, les zones sur-segmentées sont plus importantes car de nombreux éléments sont présents autour de la chevelure dans le fond de l'image. Ces résultats peuvent s'expliquer pour deux principales raisons :

- *Les éléments du fond similaire à la chevelure* : Certaines zones de l'image peuvent être marquées comme des pixels cheveux H_m car elles portent de l'information qui correspond à celle de la chevelure que l'on veut segmenter, on peut par exemple citer les vêtements.
- *Les éléments du fond non marqués* : A l'opposé, certaines zones peuvent ne pas être marquées ni pas les marqueurs cheveux H_m ou pas cheveux \bar{H}_m et comme nous avons montré dans notre étude du traitement "*Matting*" des images naturelles (section 3.3.2), ces zones peuvent être segmentées comme de la chevelure car l'influence des marqueurs cheveux est plus importante.

Les premiers résultats obtenus par cette approche assez simple sont prometteurs. Dans l'ensemble les chevelures sont bien segmentées avec peu de pertes, la proportion de sur-segmentation est plus importante. Mais dans notre traitement de fusion logique des données, une partie de l'information n'est pas exploitée et les pixels dans le masque fréquence ($F = 1$) et hors du masque de couleur ($C = 0$) ne sont pas traités. Le contrôle de ces pixels "inconnus" pourrait permettre une meilleure définition des marqueurs et une segmentation mieux guidée. Une première approche simple basée sur la définition topologique de ces zones autour du visage à partir de la détection du visage de Viola & Jones [Viola 01] a été mise en place. Mais la définition de ces zones dépendait trop du type de coupe étudiée. La prise en compte des pixels inconnus comme nouvel état possible par l'utilisation des fonctions de croyances est une approche qui nous est apparue plus adaptée.

6.3.2 Le MCT Capillaire

Dans cette section, nous présentons le formalisme des fonctions de croyances transférables appliqué à notre problématique de segmentation capillaire et plus précisément au positionnement des marqueurs du traitement "*Matting*".

6.3.2.1 Les 3 espaces de fusion de croyances

Comme dit précédemment, le principal intérêt de l'utilisation des fonctions de croyances est la possibilité de modéliser l'incertitude comme un état du système à part entière. Cette modélisation de l'incertitude correspond parfaitement avec le cadre du traitement "*Matting*" qui correspond à la zone à estimer, les 2 autres zones étant définies par les marqueurs d'information H_m (codés en blanc) et les marqueurs de fond \bar{H}_m (codés en noirs). Ainsi on peut définir l'univers Ω suivant le modèle :

$$\Omega = \{H, \bar{H}\} \quad (6.23)$$

où pour chaque pixel, H représente l'état Cheveu et \bar{H} l'état non Cheveux.

Les états E_i d'un tel système sont supposés uniques et appartiennent obligatoirement aux éléments de Ω :

$$E_i \in \Omega = \{\{\emptyset\}, \{H\}, \{\bar{H}\}, \{Un\}\} \quad \text{où} \quad \{Un\} = \{H, \bar{H}\}. \quad (6.24)$$

Notre cadre de discernement est *exhaustif*, en effet il contient forcément l'hypothèse permettant de décrire le système. Nous nous situons donc dans un monde fermé [Shafer 76]. Dans ce cas, nos FM sont normales et $m^\Omega(\emptyset) = 0$.

Pour chaque fonction de masse FM correspondant à une source (fréquence ou couleur), on associe la distribution de masse associée à chaque élément de Ω :

- $m^\Omega(H)$: qui représente la croyance pour un pixel d'être un pixel cheveu.
- $m^\Omega(\bar{H})$: qui représente la croyance pour un pixel d'être un pixel non cheveu.
- $m^\Omega(Un)$: qui représente la croyance pour un pixel soit dans l'état d'incertitude.
- $m^\Omega(\emptyset)$: qui représente la croyance pour que l'état du pixel soit inconnu.

Chaque FM vérifie l'équation (6.7) mais la répartition des masses sera définie de façon à ce qu'un pixel ne puisse pas avoir de masse à la fois sur l'état H et l'état \bar{H} en même temps. En effet vu notre approche, un pixel ne peut être à la fois un pixel cheveu et un pixel non cheveu. Ceci implique que la répartition des masses se fait toujours soit entre l'état cheveu H et l'état incertain soit entre l'état non cheveu \bar{H} et l'état incertain (*figure 6.6*). En fonction de la source étudiée, La définition de seuils a, b, c et d (*figure 6.6*) permet de contrôler la largeur de connaissances des zones de cheveu H (respectivement de non cheveu \bar{H}) grâce au seuil a ((respectivement au seuil d) ainsi que la transition entre ces zones (écart entre a et b et entre c et d). On peut aussi noter que les seuils b et c peuvent être confondus en un seul et même point.

Maintenant que nous avons décrit le formalisme général de notre modèle de fonction de croyances pour l'analyse capillaire, nous pouvons définir les FMs pour chaque descripteur caractéristique de la chevelure : la source fréquentielle (section 6.3.2.2) et la source couleur (section 6.3.2.3).

Suite à notre analyse fréquentielle (*cf chapitre 4*) et analyse couleur (*cf chapitre 5*) nous avons calculé des masques binaires (*figure 6.5*) représentant la position la plus probable des cheveux dans une image. Mais cette représentation binaire n'est pas adaptée pour le modèle de croyance. En effet dans ce cas, nous retrouvons la partition binaire pour les valeurs de seuil définies selon le modèle :

$$a = 0 \quad 0 < b \leq c < 1 \quad d = 1$$

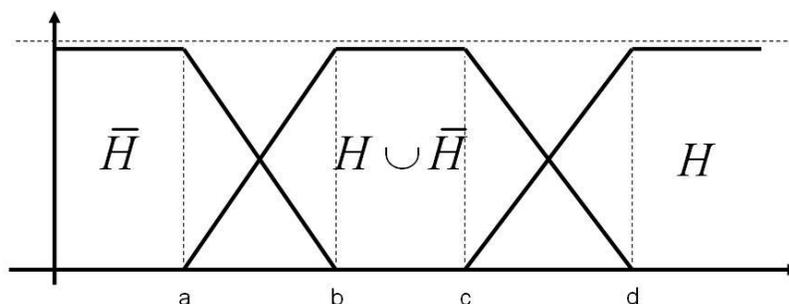


FIGURE 6.6 – **Modèle de la Distribution de masse** : Schéma de la répartition de masse d'une FM pour notre Univers $\Omega = \{H, \bar{H}\}$.

Afin d'utiliser le modèle des fonctions de croyances, les deux descripteurs seront redéfinis linéairement dans l'intervalle $[0,1]$.

6.3.2.2 Source Fréquentielle et Fonction de masse m_f

Dans notre étude fréquentielle, nous créons la carte binaire de localisation des cheveux en seuillant l'image filtrée par notre filtre gaussien. À partir de l'image filtrée, plusieurs approches sont alors possibles afin de créer une carte avec une plus grande dynamique :

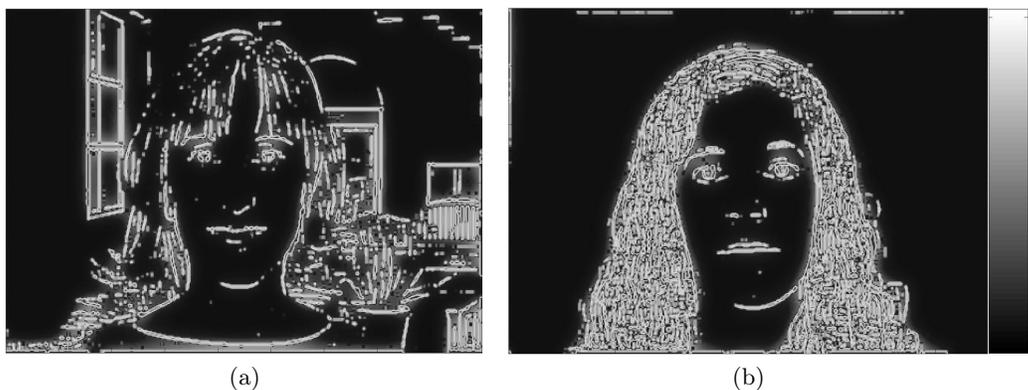


FIGURE 6.7 – **Descripteur Fréquentiel** : Représentation du descripteur fréquentiel linéaire dans l'intervalle $[0,1]$ après diffusion du masque binaire. On représente les pixels blancs pour le descripteur fréquentiel= 1 et les pixels noirs pour le descripteur fréquentiel= 0

- **L'image filtrée** : l'approche la plus intuitive serait de prendre directement l'image filtrée comme descripteur. Mais ce choix n'est pas adaptée car l'image filtrée contient trop d'information des zones de bruit.
- **L'image filtrée et seuillée** : Une deuxième approche envisagée afin de ne garder que l'information utile, est de garder l'information de l'image après seuillage. En utilisant ce descripteur nous perdons trop d'informations car elle n'est pas uniformément répartie et certaines zones de la chevelure ont des valeurs moins fortes que d'autres ce qui entraînera une mauvaise répartition de la croyance du descripteur fréquentiel.

- **L'image binaire diffuse** : à partir de la carte binaire, nous effectuons une interpolation grâce à un filtre moyenneur itératif (*figure 6.8*). Cette diffusion va permettre de lisser la carte binaire. Le principe est simple, nous filtrons l'image avec un filtre moyenneur à fenêtre glissante itérativement jusqu'à ce que l'ensemble des valeurs nulles de la carte binaire de l'image ait été filtré. Une amélioration facile pour diminuer le temps de calcul, consiste à fixer une valeur seuil pour stopper les itérations. En effet, nous pouvons admettre que pour des valeurs de pixels très faibles inférieures à $1/9$ (valeur de la diffusion d'un seul pixel par le filtre moyenneur) nous pourrions alors considérer la valeur de ces pixels comme nulle.

Les valeurs des seuils de la distribution de masse m_f du descripteur fréquentiel sont basées sur l'étude de la répartition de pixels dans un voisinage 8-connexes :

$$a = 0 \quad b = 1/9 \quad c = 2/9 \quad d = 3/9$$

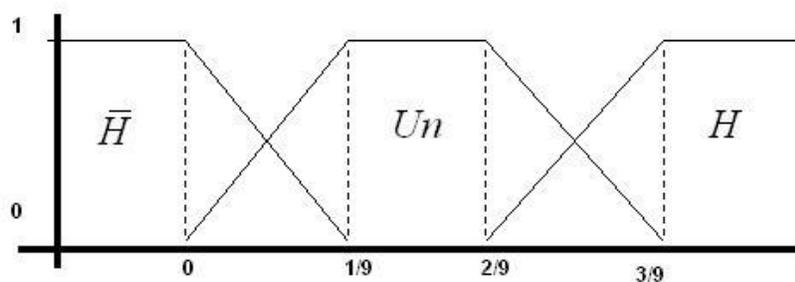


FIGURE 6.8 – Distribution Fréquentielle m_f

En considérant un pixel (respectivement 2 et 3) filtré par un filtre moyenneur de taille 3x3, nous obtenons après filtrage une valeur de pixel de $1/9$ (respectivement $2/9$ et $3/9$). Ainsi pour un unique pixel dans un voisinage 3x3 que l'on peut considérer comme un artefact de segmentation, on obtient la valeur de $1/9$ qui représente la plus petite valeur pour émettre un doute sur la possibilité de trouver de l'information capillaire pertinente qui permet de définir le seuil b . Le seuil c est fixé à $2/9$ et le seuil d à $3/9$. Nous admettons que pour 3 pixels dans un voisinage 8-connexe que nous avons suffisamment d'information fréquentielle pour définir avec certitude la limite de la zone H . Le seuil a est fixé à zéro car il implique qu'il n'y avait aucun pixel dans le voisinage et donc on peut définir avec certitude la limite de la zone \bar{H}

6.3.2.3 Source Couleur et Fonction de masse m_c

La modélisation du descripteur couleur est plus simple que pour le descripteur fréquence. En effet la distribution en gaussienne du modèle couleur permet une définition linéaire sans avoir à le redéfinir. Les valeurs des seuils de la distribution de masse m_c du descripteur couleur sont basées sur l'étude de la distribution gaussienne du modèle couleur.



FIGURE 6.9 – **Descripteur Couleur** : Représentation du descripteur fréquentiel linéaire dans l'intervalle $[0,1]$. On représente les pixels noirs pour le descripteur couleur= 1 et les pixels blancs pour le descripteur couleur= 0

Ainsi en faisant le parallèle entre le pourcentage d'appartenance d'appartenir à la distribution couleur et le pourcentage d'appartenir au modèle couleur, nous pouvons facilement fixer les seuils en fonction des valeurs de l'écart type de la distribution gaussienne selon le modèle :

$$a = 4\sigma \quad b = 3\sigma \quad c = 2\sigma \quad d = \sigma$$

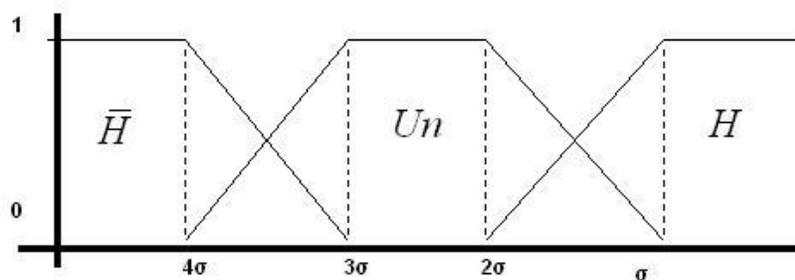


FIGURE 6.10 – **Distribution Couleur** m_c

6.3.2.4 DRC Capillaire

m_c/m_f	H	\bar{H}	Un
H	H	Un	Un
\bar{H}	Un	\bar{H}	Un
Un	Un	Un	Un

TABLE 6.4 – Table des règles de Combinaison Disjonctive pour le modèle Capillaire

L'application de la DRC joue un rôle de *généralisation* qui favorise l'état d'incertitude Un et qui peut amener à une FM vide (section 6.2.1.3) en particulier pour les sources d'information qui ne sont pas fiables.

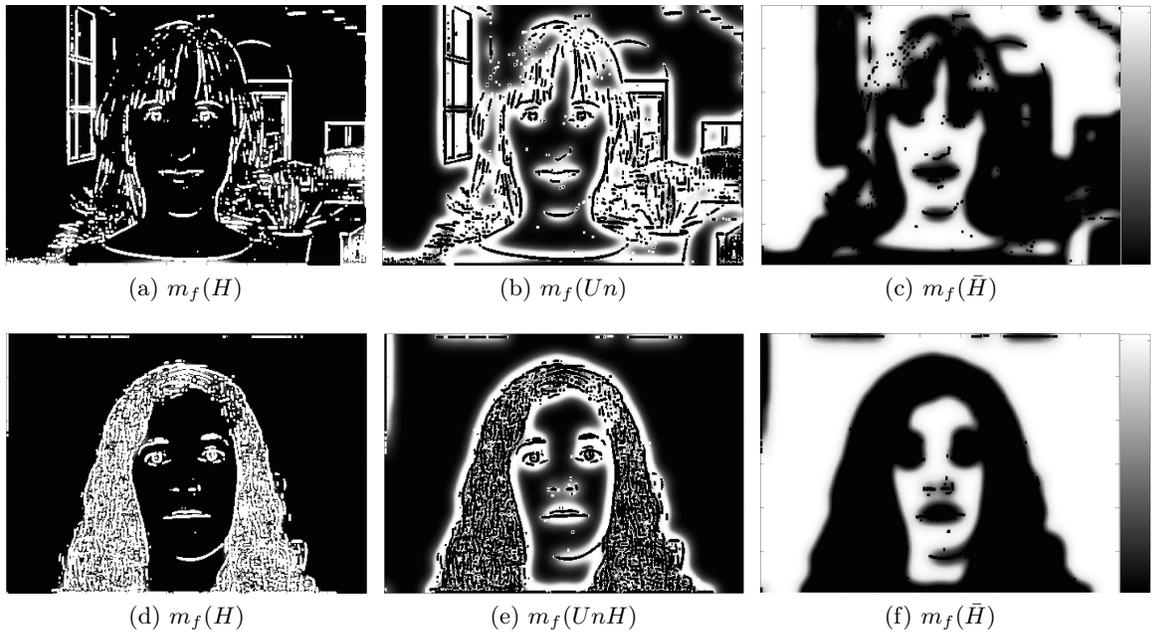


FIGURE 6.11 – **Fonction de masse m_f** : Représentation de la fonction de masse fréquentielle pour chaque élément de Ω . On représente les pixels blancs pour $m_f = 1$ et les pixels noirs pour $m_f = 0$

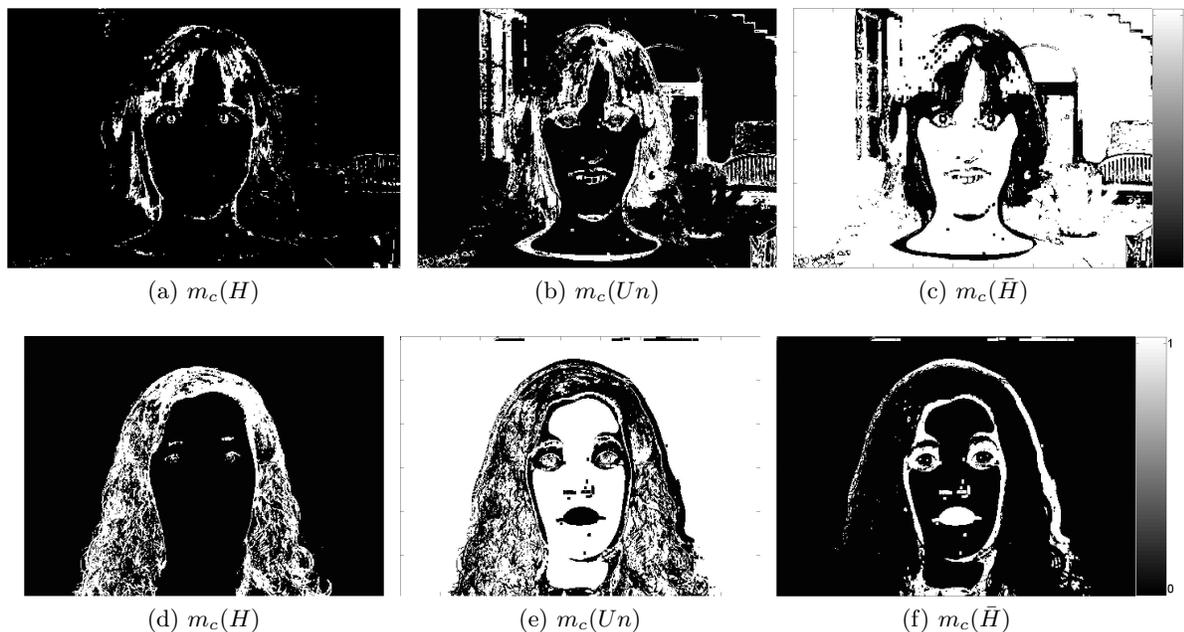


FIGURE 6.12 – **Fonction de masse m_c** : Représentation de la fonction de masse couleur pour chaque élément de Ω . On représente les pixels blancs pour $m_c = 1$ et les pixels noirs pour $m_c = 0$

La DRC privilégie la zone d'incertitude Un ce qui agrandit la zone où l'alpha matte devra être estimé tout en réduisant les zones de définition des marqueurs H_m et \bar{H}_m . Comme notre univers Ω est binaire, l'application de la DRC revient presque à une fusion logique.

6.3.2.5 CRC Capillaire

Contrairement à la DRC, l'application de la CRC joue un rôle de *spécialisation* qui favorise les états H et \bar{H} :

m_c/m_f	H	\bar{H}	Un
H	H	\emptyset	H
\bar{H}	\emptyset	\bar{H}	\bar{H}
Un	H	\bar{H}	Un

TABLE 6.5 – Table de règles de Combinaison Conjonctive pour le modèle Capillaire

$$\begin{aligned}
m_{fc}(\emptyset) &= m_f(H) \cdot m_c(\bar{H}) + m_f(\bar{H}) \cdot m_c(H) \\
m_{fc}(H) &= m_f(H) \cdot m_c(H) + m_f(Un) \cdot m_c(H) + m_f(H) \cdot m_c(Un) \\
m_{fc}(\bar{H}) &= m_f(\bar{H}) \cdot m_c(\bar{H}) + m_f(Un) \cdot m_c(\bar{H}) + m_f(\bar{H}) \cdot m_c(Un) \\
m_{fc}(Un) &= m_f(Un) \cdot m_c(Un)
\end{aligned} \tag{6.25}$$

Par la combinaison conjonctive, nous définissons une nouvelle FM qui n'est plus normale, en effet nous avons une masse sur l'ensemble vide qui n'est pas nulle ($m_{fc}(\emptyset) = m_f(H) \cdot m_c(\bar{H}) + m_f(\bar{H}) \cdot m_c(H)$) (figures 6.13a 6.13d). Nous avons créé une source de conflit tel que nous l'avons défini dans la section 6.2.4. Nous traitons ce conflit par la règle de Yager [Yager 87] qui transfère le conflit $m_{fc}(\emptyset)$ sur le doute total $m_{fc}(Un)$ et nous redéfinissons une FM normale ($m_{fc}(\emptyset) = 0$) selon le modèle suivant :

$$\begin{aligned}
m_{fc}(\emptyset) &= 0 \\
m_{fc}(H) &= m_f(H) \cdot m_c(H) + m_f(Un) \cdot m_c(H) + m_f(H) \cdot m_c(Un) \\
m_{fc}(\bar{H}) &= m_f(\bar{H}) \cdot m_c(\bar{H}) + m_f(Un) \cdot m_c(\bar{H}) + m_f(\bar{H}) \cdot m_c(Un) \\
m_{fc}(Un) &= m_{fc}(Un) + m_{fc}(\emptyset) \\
&= m_f(Un) \cdot m_c(Un) + m_f(H) \cdot m_c(\bar{H}) + m_f(\bar{H}) \cdot m_c(H)
\end{aligned} \tag{6.26}$$

6.3.2.6 Prise de Décision Capillaire

Normalement la prise de décision est prise à partir de la distribution de probabilité pignistique (section 6.2.3.1). Cependant, notre application pour le traitement "Matting" se base sur 2 hypothèses qui permettent de faciliter la prise de décision :

- *L'univers Ω* : notre univers d'étude est binaire et ne comprend que les états cheveux H ou non cheveux \bar{H} .
- *L'état incertain Un* : l'état d'incertitude est considéré comme un état possible pour un pixel.

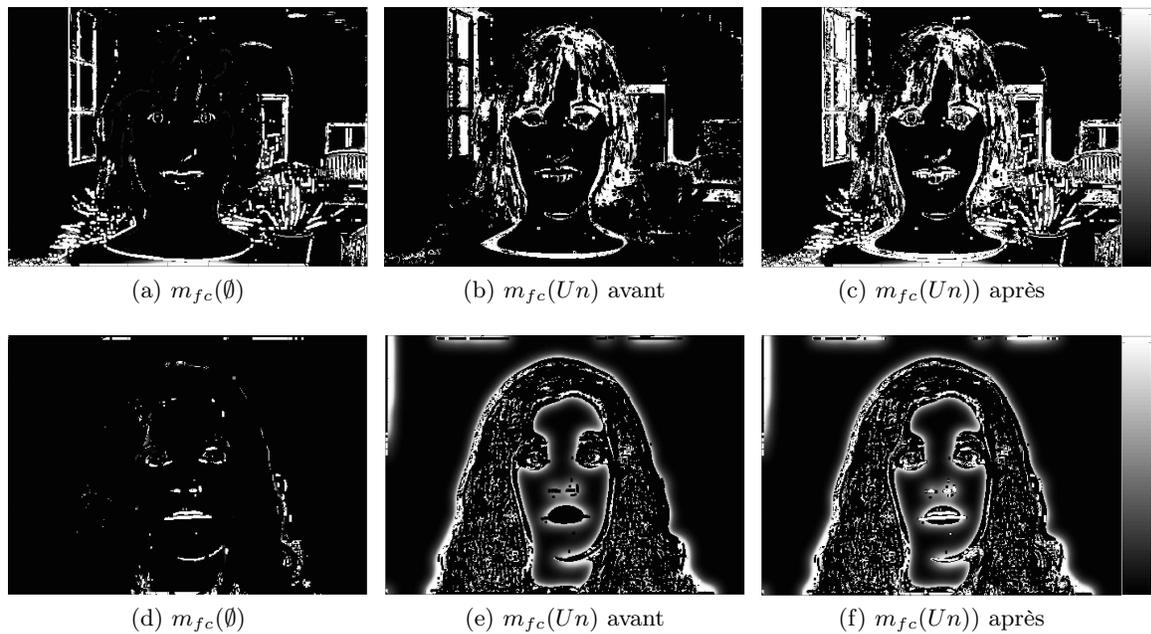


FIGURE 6.13 – **Gestion du conflit** : Représentation du transfert du conflit $m_{fc}(\emptyset)$ sur le doute total $m_{fc}(Un)$ par la règle de Yager. On représente les pixels blancs pour $m_{fc} = 1$ et les pixels noirs pour $m_{fc} = 0$

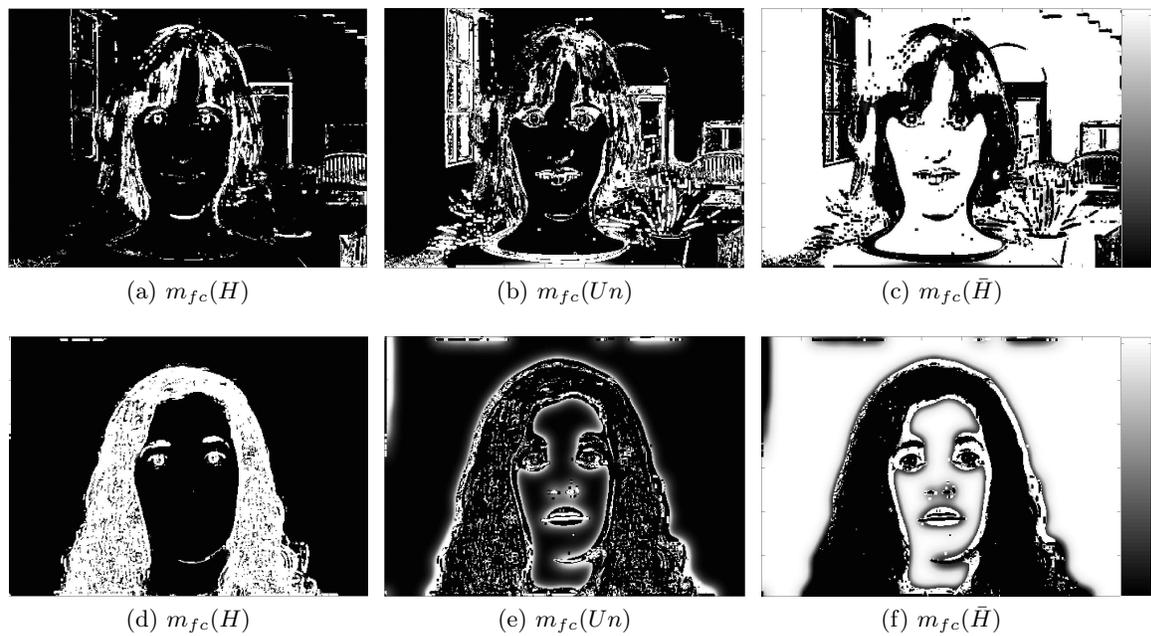


FIGURE 6.14 – **Fonction de masse** m_{fc} : Représentation de la fonction de masse issue de la CRC pour chaque élément de Ω . On représente les pixels blancs pour $m_{fc} = 1$ et les pixels noirs pour $m_{fc} = 0$

Ainsi la prise de décision de l'état d'un pixel est déterminé par la distribution des fonctions de masses pour chaque état de l'univers Ω :

$$\omega_r = \arg \max_{\omega_k \in \Omega} m^\Omega(\omega_k) \quad (6.27)$$

6.3.2.7 Résultats fonctions de croyances

Nous évaluons notre approche de fusion par le modèle des fonctions de croyances pour 200 images de la base maquillage et 160 images de la base Caltech. Nous comparons les résultats avec ceux obtenus par fusion logique pour évaluer notre nouvelle approche de fusion de données. Les résultats sont calculés sur les masques capillaires binaires obtenus après segmentation par le traitement "*Matting*" initialisé par les marqueurs définis par la fusion par fonctions de croyances.

	T.F.R.	σ (T.F.R.)	T.F.A.	σ (T.F.A.)
Fusion Logique (base maquillage)	9.1%	12.7%	14.2%	15.4%
Fusion Logique (base Caltech)	6.7%	16.3%	40.6%	21.6%
Fonctions de croyances (base maquillage)	8.86%	13.4%	14%	18.2%
Fonctions de croyances (base Caltech)	7.32%	15.4%	40%	26.67%

TABLE 6.6 – Taux de faux rejet (T.F.R.) et de fausse acceptance (T.F.A.) pour la segmentation capillaire par fonctions de croyances. Résultats obtenus pour 200 images de la base maquillage et 160 images de la base Caltech

Les deux approches donnent des résultats quasiment identiques. Nous obtenons un taux de faux rejet de 9.1% et de 6.7% pour la fusion logique et un taux de faux rejet de 8.86% et de 7.32% pour les fonctions de croyances. Pour le taux de fausse acceptance, nous observons un taux de 14.2% et de 40.6% pour la fusion logique un taux de 14% et de 40% pour les fonctions de croyances. Même si les deux approches sont différentes et procurent des marqueurs H_m et \bar{H}_m différents, l'estimation de l'alpha matte permet d'obtenir des masques capillaires similaires.

Même si l'utilisation des fonctions de croyances n'a pas permis l'amélioration de la segmentation capillaire, ces résultats montrent que cette approche est néanmoins adaptée pour la définition des marqueurs du traitement "*Matting*". Pour améliorer nos résultats, nous allons intégrer le 3^{ème} paramètre caractéristique de la chevelure : sa localisation.

6.3.3 Fonction d'affaiblissement

Dans certaines applications, il est possible de quantifier la fiabilité des sources ce qui permet d'appliquer une règle de combinaison moins conservatrice dans le cas de sources peu fiables. La prise en compte de la fiabilité dans le cadre des fonctions de croyances porte le nom d'affaiblissement car le processus consiste à pondérer la masse des éléments d'une FM. Les premiers travaux sur l'affaiblissement dans le cadre des fonctions de croyances ont été développés par Shafer [Shafer 76], Smets [Smets 93] et généralisés par Mercier et Denoeux [Mercier 08].

6.3.3.1 Formalisme

L'affaiblissement de Shafer d'une FM m est défini comme la pondération de chaque masse de la distribution par un coefficient $(1 - \alpha)$ appelé fiabilité et où $\alpha \in [0, 1]$ est le taux d'affaiblissement.

$$\begin{aligned} \alpha m^\Omega(B) &= (1 - \alpha) \cdot m^\Omega(B), \forall B \subset \Omega \\ \alpha m^\Omega(\Omega) &= (1 - \alpha) \cdot m^\Omega(\Omega) + \alpha \end{aligned} \quad (6.28)$$

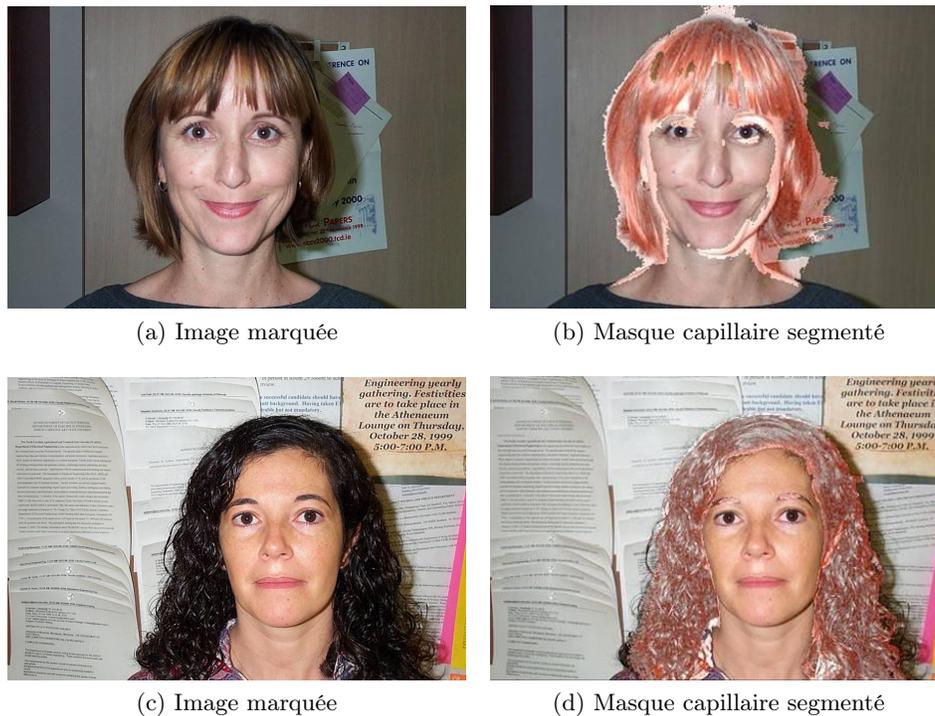


FIGURE 6.15 – **Segmentation par fonctions de croyances** : Exemple de segmentations réussies après traitement "Matting" initialisé par les marqueurs obtenus par fonctions de croyances.

Plus la source est fiable, moins la FM est modifiée et inversement plus la fiabilité diminue, plus la FM issue de l'affaiblissement tend vers la FM vide.

Pour prendre en compte l'information de localisation de la chevelure, nous nous sommes basés sur une constatation assez simple : Même si une chevelure peut prendre beaucoup de formes possibles, elle demeure a priori autour du visage. C'est sur cette hypothèse que nous nous sommes appuyés pour modéliser la fiabilité de nos sources.

Nous avons vu dans la littérature que généralement l'approche utilisée consistait en la modélisation d'une distribution de probabilité calculée simplement en comptant pour chaque pixel le nombre de pixels appartenant à un masque capillaire à partir d'une base d'apprentissage [Lee 08] [Wang 09]. Une telle définition de la carte de localisation montre rapidement des limites. La première est la constitution d'une telle distribution qui nécessite une base d'entraînement comprenant une grande variété de coiffures possibles. Cette approche vise à maximiser la ressemblance entre toutes les coiffures. Au cours de notre étude nous avons constaté que pour la segmentation capillaire, il valait mieux suivre une approche pour minimiser les différences. Pour cela nous avons mis une approche plus globale de la localisation d'une chevelure.

À partir de la segmentation du masque de peau que nous avons mise en place à la section 5.3, nous créons 2 cartes de distances :

- Une carte de distance à l'extérieur du visage D_{out} : Plus un pixel sera loin du visage moins il aura de chance d'appartenir au masque capillaire.

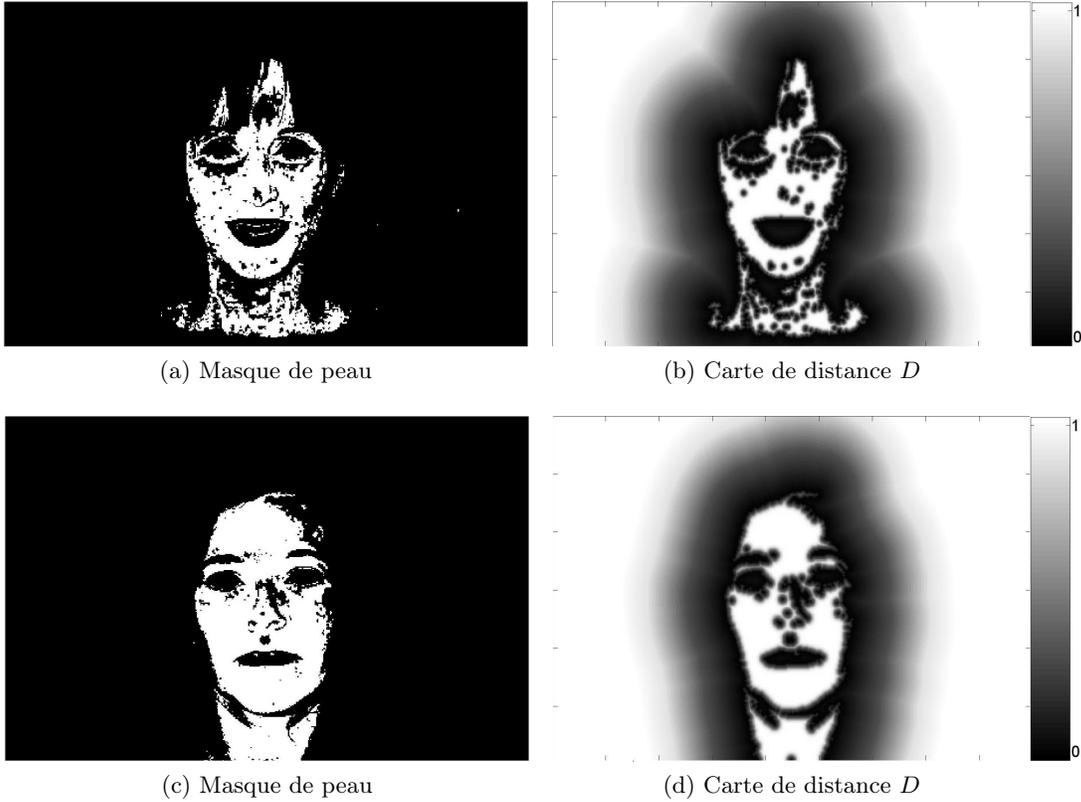


FIGURE 6.16 – **Carte de Distance D** : Représentation de la carte de distance en fonction du masque de peau. On représente les pixels blancs pour $D = 1$ et les pixels noirs pour $D = 0$.

- Une carte de distance à l’intérieur du visage D_{in} : Plus un pixel sera au centre du visage moins il aura de chance d’appartenir au masque capillaire. Même si en général une chevelure est autour d’un visage, nous avons créé cette carte pour prendre en considération les mèches qui peuvent recouvrir une partie du visage ainsi que pour prendre en compte le problème de sous-segmentation des racines que nous avons vu pour la fusion logique (section 6.3.1).

Au final nous sommes les 2 cartes de distances pour définir la carte de distance $D = D_{in} + D_{out}$ (figure 6.16).

Nous intégrons la fonction d’affaiblissement à notre MCT capillaire en se basant sur la modélisation de Denoex [Mercier 08] défini par le système d’équation suivant :

$$\begin{aligned}
 m_D(\emptyset) &= 0 \\
 m_D(H) &= (1 - D) \cdot m_{fc}(H) \\
 m_D(\bar{H}) &= D \cdot m_{fc}(\bar{H}) \\
 m_D(Un) &= m_{fc}(Un) + D \cdot m_{fc}(H) + (1 - D) \cdot m_{fc}(\bar{H})
 \end{aligned}
 \tag{6.29}$$

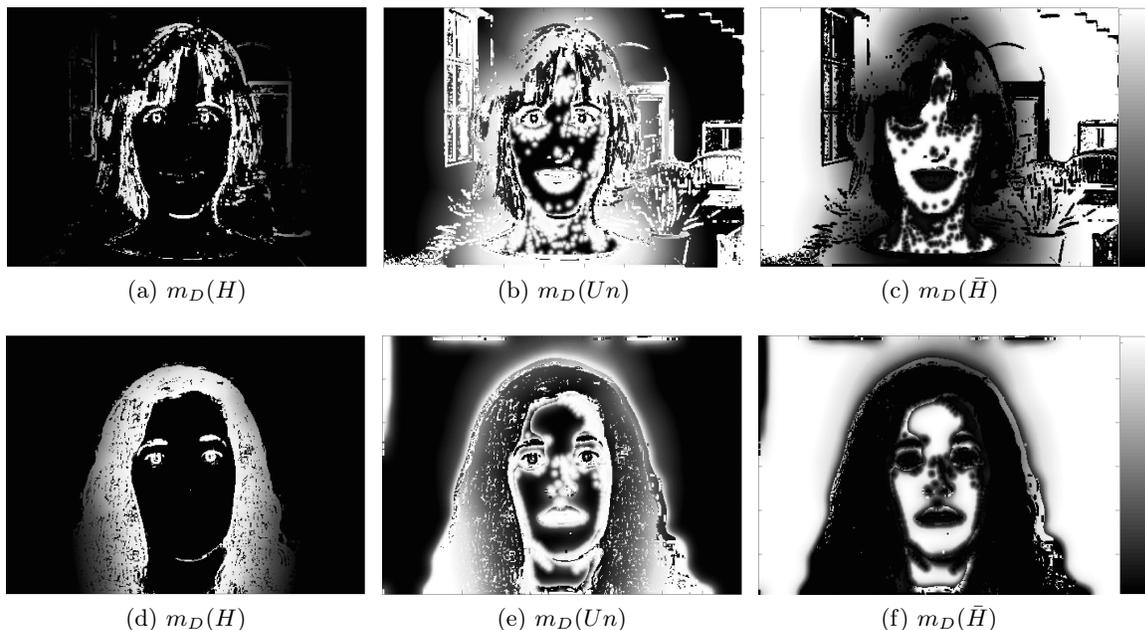


FIGURE 6.17 – **Fonction de masse m_D** : Représentation de la fonction de masse après affaiblissement des marqueurs H_m et \bar{H}_m pour chaque élément de Ω . On représente les pixels blancs pour $m_D = 1$ et les pixels noirs pour $m_D = 0$

6.3.3.2 Résultats

Nous évaluons l'apport du paramètre capillaire de la localisation par l'ajout d'une fonction d'affaiblissement à notre MCT capillaire pour 160 images de la base Caltech. Nous avons testé notre algorithme uniquement sur la base Caltech car la base maquillage étant prise pour des fonds uniformes, elle n'apportait par d'information exploitable pour évaluer l'amélioration de nos résultats. Les résultats sont calculés sur les masques capillaires binaires obtenus après segmentation par le traitement "*Matting*" initialisé par les marqueurs définis par fonctions de croyances affaiblies. Nous comparons nos résultats pour 2 cartes de distances :

- *La carte de visage uniquement (1)*
- *La carte de visage avec cou (2)* : la prise en compte du cou va permettre de mieux traiter les chevelures aux cheveux longs.

	T.F.R.	σ (T.F.R.)	T.F.A.	σ (T.F.A.)
Fonctions de croyances	7.32%	15.4%	40%	26.67%
Fonctions de croyances affaiblies (1)	2.68%	4.34%	37.63%	20.96%
Fonctions de croyances affaiblies (2)	4.6%	5.4%	30.7%	15.4%

TABLE 6.7 – Taux de faux rejet (T.F.R.) et de fausse acceptance (T.F.A.) pour la segmentation capillaire par fonctions de croyances avec affaiblissement. Résultats obtenus pour 160 images de la base Caltech

Nous obtenons un très faible taux de faux rejet de seulement 2.68% et de 4.6% pour les 2 cartes de distances. Non seulement nous avons de meilleurs résultats que pour les fonctions

de croyances sans affaiblissement qui donnaient un taux de faux rejet de 7.32% mais surtout nous avons amélioré la robustesse de notre algorithme avec une baisse importante de l'écart type moyen. Ces résultats montrent que notre algorithme de segmentation capillaire est efficace et que pour l'ensemble des images traitées nous trouvons la chevelure dans sa "totalité". Ces résultats peuvent s'expliquer pour plusieurs raisons :

- *Le retrait des artefacts* : la principale amélioration apportée par la fonction d'affaiblissement est le nettoyage des pixels similaires à la chevelure se trouvant trop loin du visage. Nous retirons ainsi de nombreux pixels artefacts.
- *Baisse de l'influence des pixels non cheveux en bordures du masque capillaire* : nous augmentons la zone d'incertitude Un autour du visage et nous diminuons à l'opposé l'influence des marqueurs non cheveux H_m qui étaient trop près de la chevelure. Cette redistribution des marqueurs améliorent l'estimation de l'alpha matte.
- *Correction de la mauvaise définition des cartes couleur et fréquentielle* : En réduisant l'influence des marqueurs non cheveux \bar{H}_m proche du visage, nous corrigeons la mauvaise définition des cartes fréquentielles et couleurs. Nous éliminons en particulier les 'trous' de la carte fréquentielle dues généralement aux mèches contenant des reflets spéculaires qui saturent les valeurs des pixels. Nous éliminons aussi les zones de la carte couleur mal marquées à cause d'une trop grande variation de teinte au sein d'une chevelure.

Pour le taux de fausse acceptance, les résultats sont très peu améliorés avec un taux de 37.63% pour l'affaiblissement (1) contre 40% sans. Les résultats pour la fonction de distance avec le cou (2) sont légèrement meilleurs avec un taux de fausse acceptance de seulement 30.7%. L'amélioration obtenue est faible et nous sur-segmentons encore beaucoup d'images.

Même si nous avons un taux de fausse acceptance élevé, nous avons déplacé la zone de mauvaise estimation de l'alpha matte. Initialement elle pouvait se situer dans toute l'image, maintenant elle se trouve généralement dans une couronne autour du visage. Cette nouvelle localisation s'explique par la définition de notre fonction d'affaiblissement. En effet pour un certain intervalle de la carte de distance, nous diminuons en même temps l'influence des marqueurs cheveux et non cheveux \bar{H}_m et donc nous augmentons l'influence de la zone d'incertitude Un où l'alpha matte reste à estimer. Nous avons donc aussi amélioré cette composante de notre algorithme de segmentation. En effet maintenant que la sur-segmentation de la chevelure est contrôlée, nous pouvons envisager une seconde boucle de traitement pour affiner notre définition du masque capillaire.

Nous observons que l'affaiblissement simultané des marqueurs cheveux H_m et non cheveux \bar{H}_m entraîne une perte d'information. Nous avons modifié notre affaiblissement pour ne pondérer que les marqueurs H_m . Ainsi nous allons réduire la sur-segmentation en limitant l'influence des marqueurs H_m pour les pixels loin du visage tout en gardant l'influence des marqueurs non cheveux \bar{H}_m en périphérie du visage.

Nous définissons le modèle simplifié d'affaiblissement toujours basé sur les travaux de Denoex [Mercier 08] selon le système d'équations :

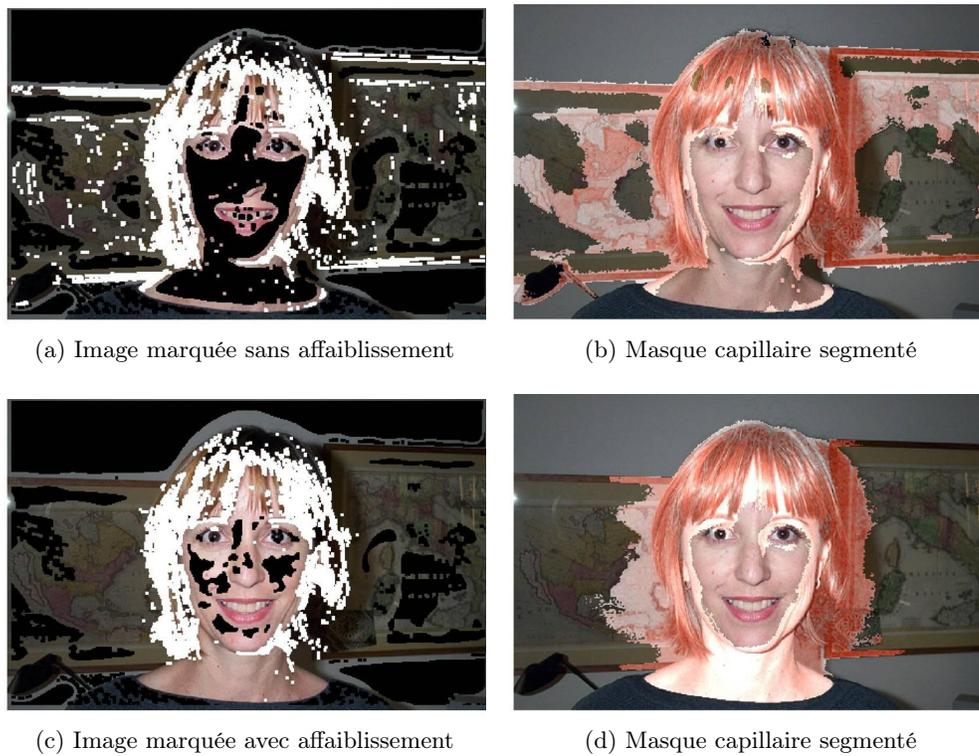


FIGURE 6.18 – Segmentation par fonctions de croyances affaiblies : Comparaison de marquage et de masque capillaire segmentée entre une fusion de croyances affaiblies ou non.

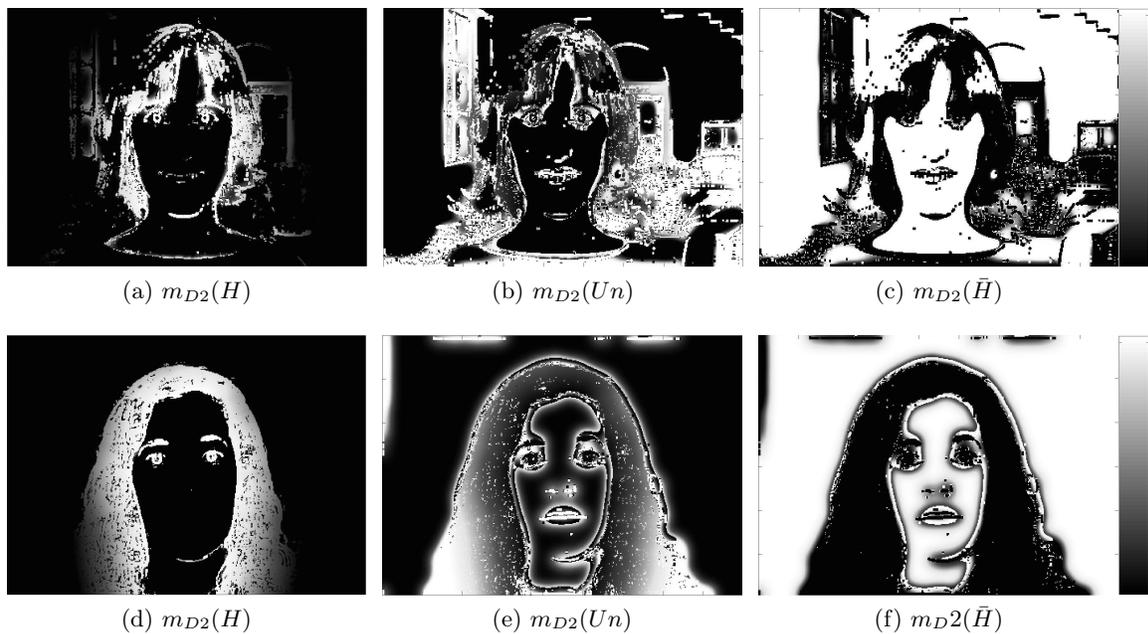


FIGURE 6.19 – Fonction de masse m_{D2} : Représentation de la fonction de masse après affaiblissement des marqueurs H_m pour chaque élément de Ω . On représente les pixels blancs pour $m_{D2} = 1$ et les pixels noirs pour $m_{D2} = 0$.

$$\begin{aligned}
m_{D2}(\emptyset) &= 0 \\
m_{D2}(H) &= (1 - D) \cdot m_{fc}(H) \\
m_{D2}(\bar{H}) &= m_{fc}(\bar{H}) \\
m_{D2}(Un) &= m_{fc}(Un) + D \cdot m_{fc}(H)
\end{aligned} \tag{6.30}$$

Nous évaluons cette alternative (3) en la comparant avec les 2 autres fonctions d'affaiblissement mises en place pour 160 images de la base Caltech :

	T.F.R.	σ (T.F.R.)	T.F.A.	σ (T.F.A.)
Fonctions de croyances affaiblies (1)	2.68%	4.34%	37.63%	20.96%
Fonctions de croyances affaiblies (2)	4.6%	5.4%	30.7%	15.4%
Fonctions de croyances affaiblies (3)	13.6%	13%	12.9%	9.1%

TABLE 6.8 – Taux de faux rejet (T.F.R.) et de fausse acceptance (T.F.A.) pour la segmentation capillaire par fonctions de croyances avec affaiblissement. Résultats obtenus pour 160 images de la base Caltech

Pour cette nouvelle approche d'affaiblissement nous obtenons un taux de faux rejet assez élevé de 13.6%. Nous segmentons moins bien la chevelure en particulier celles aux cheveux longs. Nous pouvons expliquer ces résultats par le fait que nous diminuons l'influence des marqueurs cheveux H_m sans modifier celle des marqueurs non cheveux \bar{H}_m pour les pixels de distance moyenne. Par cette approche, nous améliorons toujours la segmentation par le retrait des pixels artefacts loin du visage. Par contre nous diminuons la correction apportée sur les marqueurs \bar{H}_m résultant des faiblesses de nos cartes fréquentielles et couleurs.

En contre partie par cette approche nous avons considérablement amélioré la sur-segmentation de l'image en descendant le taux de fausse acceptance à seulement une valeur de 12.9%. Ce taux est vraiment bas par rapport aux taux de 37.63% et 30.7% obtenus pour les 2 premières cartes d'affaiblissement.

Ces résultats sont d'autant plus intéressants que nous avons ainsi créé deux méthodes d'affaiblissement qui permettent de contrôler la segmentation du masque capillaire. En fonction des résultats voulus, nous pourrions choisir une méthode plutôt qu'une autre en fonction du résultat souhaité :

- *Affaiblissement des marqueurs H_m et \bar{H}_m* : Bonne segmentation de la chevelure mais forte possibilité de sur-segmentation mais limitée dans une couronne autour du visage.
- *Affaiblissement uniquement des marqueurs \bar{H}_m* : Assez bonne segmentation avec néanmoins possibilité de sous-segmentation en particulier pour les cheveux longs. Nous obtenons par contre très peu de sur-segmentation de la chevelure.

Nous avons envisagé une approche combinée pour obtenir la meilleure segmentation possible en utilisant les avantages des deux méthodes. Une première boucle pour la définition générale du masque capillaire par la première approche d'affaiblissement puis une optimisation du masque par la deuxième approche. Les premiers résultats obtenus sans optimisation du traitement entre les deux boucles ont montré rapidement les limites d'une telle approche. En effet, quelle que soit la zone d'étude, un marquage particulier mènent sensiblement toujours à la même estimation de

l'alpha matte.

6.3.3.3 Pondération de la fusion Logique

Pour finir notre étude et valider l'apport du modèle de fonctions de croyances, nous avons exploré une dernière piste. Pour ajouter l'information de localisation de la chevelure et faire le parallèle avec la fonction d'affaiblissement, nous avons pondéré les marqueurs obtenus par notre approche par fusion logique en utilisant la même carte de distance D . Les valeurs α du marqueur cheveux H_m (respectivement non cheveux \bar{H}_m) sont pondérées de $\alpha = 1$ (respectivement $\alpha = 0$) à l'intervalle $[0.5, 1]$ (respectivement $[0, 0.5]$) (figures 6.20a 6.20c). Les résultats sont calculés sur les masques capillaires binaires obtenus après segmentation par le traitement "Matting" initialisé par les marqueurs pondérés par la carte de distance D .

	T.F.R.	$\sigma(\text{T.F.R.})$	T.F.A.	$\sigma(\text{T.F.A.})$
Fusion logique pondérée	22.73%	20.32%	45.38%	9.59%

TABLE 6.9 – Taux de faux rejet (T.F.R.) et de fausse acceptance (T.F.A.) pour la segmentation capillaire par pondération des marqueurs obtenus par fusion logique. Résultats obtenus pour 160 images de la base Caltech

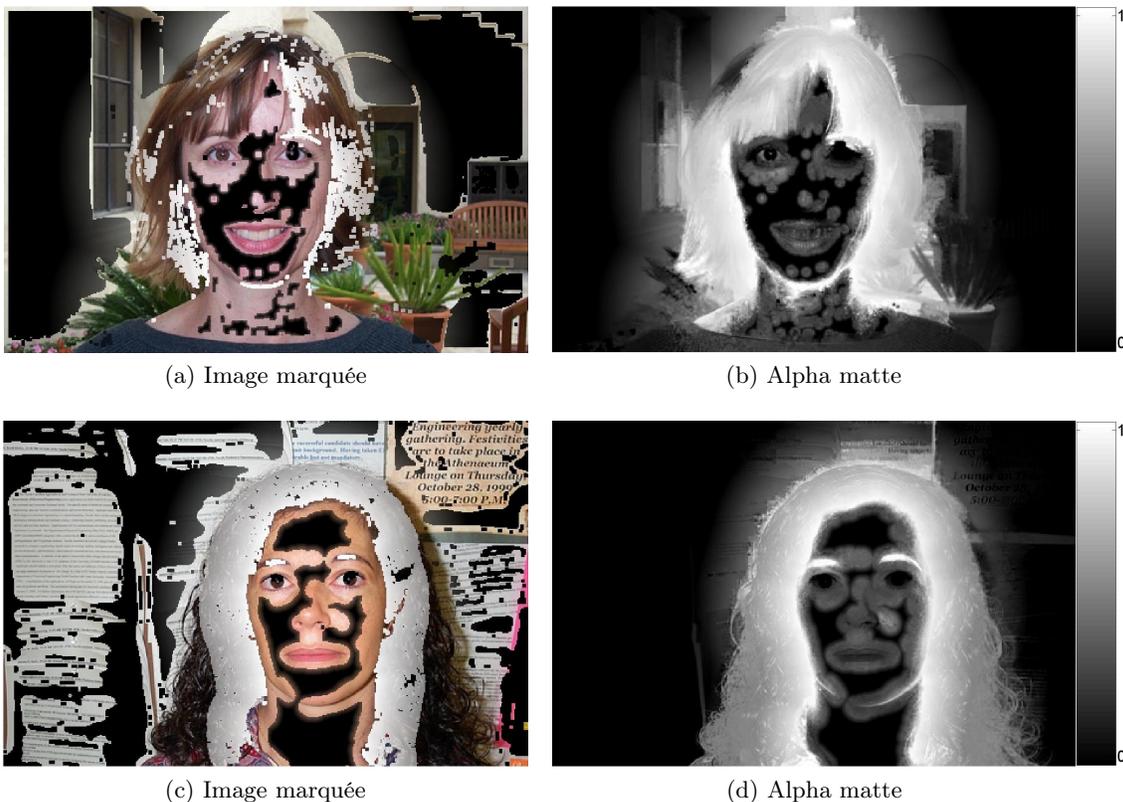


FIGURE 6.20 – Pondération de la fusion Logique : Exemple d'alpha matte estimé après traitement "Matting" initialisé par des marqueurs pondérés par une carte de distance au visage

Nous obtenons de mauvais résultats avec un taux de faux rejet de 22.73% et un taux de

fausse acceptance de 45.38%. Cette approche n'est pas adaptée et ne permet pas d'obtenir de bonnes segmentations car la pondération est effectuée sur l'influence des marqueurs et non pas sur leurs positions. Les marqueurs vont toujours être positionnés pour les mêmes pixels mais c'est leur valeur de α qui a été modifiée. Par définition, la valeur α définit le plan de mélange liant les deux plans \mathbf{F} et \mathbf{B} d'une image composite. Cette composante peut prendre des valeurs dans l'intervalle $[0, 1]$ avec un pixel transparent pour $\alpha = 0$ et un pixel opaque pour $\alpha = 1$. Ainsi un pixel qui a une valeur de $\alpha = 0.5$ appartient aux deux plans. En pondérant les valeurs des marqueurs par la carte de distance, on modifie l'information qu'ils portent et à la limite de la fonction de distance on définit les marqueurs pour les 2 états. Par cette approche en diminuant le marquage cheveux on augmente le marquage non cheveux et inversement. Finalement après estimation nous obtenons un alpha matte très diffus avec des valeurs de α autour de 0.5 (*figures 6.20b 6.20d*).

6.4 Conclusion

Dans ce chapitre nous avons introduit le procédé de fusion de données pour combiner l'information des masques binaires de fréquences et de couleurs. Nous avons présenté en détail l'approche par fonctions de croyances transférables. Puis nous avons proposé un ensemble d'algorithmes pour segmenter une chevelure dans une image. Dans un premier temps nous avons mis en place une approche par fusion logique pour évaluer la fiabilité des deux paramètres capillaires pour la localisation d'une chevelure. Nous avons obtenu des premiers résultats intéressants. Pour optimiser la définition des marqueurs du traitement "*Matting*", nous avons ensuite mis en place une approche fusion par fonctions de croyances. Nous avons confirmé nos résultats en obtenant des segmentations proches de celles obtenues par fusion logique. Afin de prendre en compte le 3^{ème} paramètre capillaire, nous avons intégré une fonction d'affaiblissement basé sur une carte de distance au visage. Nous avons alors utilisé deux approches : une première où on effectuait l'affaiblissement sur les deux types de marqueurs cheveux H_m et non cheveux \bar{H}_m , la seconde où on effectuait l'affaiblissement seulement sur les marqueurs cheveux H_m . Nous avons obtenu ainsi deux algorithmes donnant de bons résultats et permettant de contrôler la sur-segmentation de l'image. Pour finir nous avons comparé l'apport des fonctions d'affaiblissement avec une pondération des marqueurs obtenus par fusion logique.

Notre algorithme de segmentation capillaire donne de bons résultats. La segmentation du masque capillaire est généralement réussie avec dans la majorité des cas l'ensemble de la chevelure qui est segmentée. Bien que notre approche ait tendance à sur-segmenter l'image, nos propositions d'optimisations de notre algorithme permettent de contenir le nombre et la répartition de ces pixels. Nous présentons un algorithme qui est stable pour une grande variété de coiffures. Nous présentons aussi une technique qui est robuste pour différentes conditions d'acquisitions que ce soit des variations d'illuminations ou des conditions de prise de vue avec des fonds chargés.

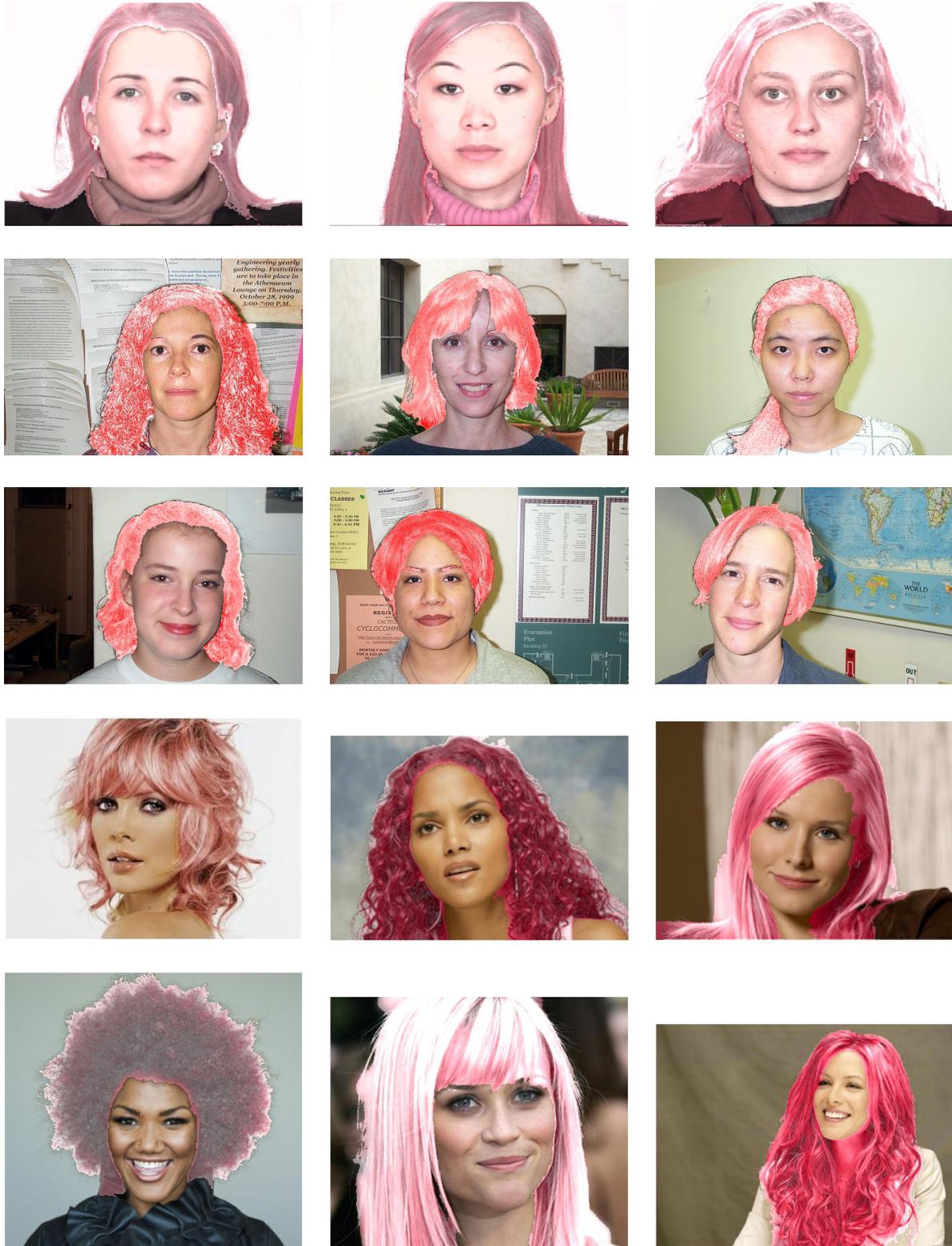
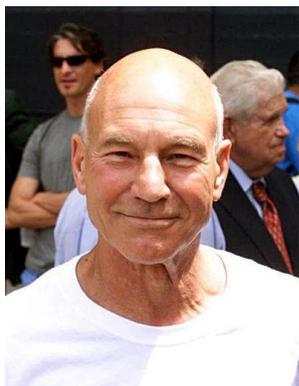


FIGURE 6.21 – **Segmentation Capillaire - Résultats** : Exemples de segmentations du masque capillaire avec coloration non naturelle pour faire ressortir les zones segmentées.

7 Caractérisation du masque capillaire



Sommaire

7.1	Caractérisation de la Texture	155
7.2	Caractérisation de la Teinte	158
7.3	Coloration du masque capillaire	160
7.4	Caractérisation de la Forme	163
7.5	Conclusion	169

Dans les chapitres précédents, nous avons présenté nos travaux sur l'algorithme de segmentation du masque capillaire que nous avons mis en place. Dans ce chapitre, nous proposerons une première étude sur la caractérisation du masque capillaire. Nous proposerons des pistes pour la définition d'outils basés sur les 3 paramètres caractéristiques de la chevelure : la texture, la couleur et la position de la chevelure autour du visage.

Nous avons limité nos travaux à une simple approche par traitement d'image afin de fournir des outils de description de la chevelure où nous avons volontairement ignoré l'aspect cognitif car nous ne possédons pas l'expertise nécessaire pour définir des paramètres cognitifs cohérents pour la description de la chevelure.

Dans un premier temps, nous présenterons notre analyse de la texture capillaire par la définition d'un masque vectoriel représentant la direction des cheveux à l'intérieur du masque capillaire (section 7.1). Puis nous présenterons notre analyse de la teinte de la chevelure pour la définition d'un classifieur couleur (section 7.2). Enfin nous présenterons notre analyse de la position de la chevelure par la création d'une signature basée sur la répartition des pixels cheveux autour du visage (section 7.4).

7.1 Caractérisation de la Texture

Dans le chapitre 4, nous avons déterminé la position de la chevelure par l'étude de sa texture à partir d'une analyse fréquentielle. Nous avons recherché la bande de fréquence qui correspondait le mieux à celle de la chevelure, pour cela nous avons commencé notre étude par l'utilisation de filtres orientés GLOP. Finalement nous sommes revenu à une approche par un filtre isotrope pour obtenir une carte de localisation fréquentielle multi-directionnelle car nous cherchons à caractériser la chevelure quel que soit son orientation. En utilisant cette information directionnelle et en se basant sur les travaux de synthèse de chevelure de Chen *et al* [Chen 06], nous avons créé un masque représentant les directions privilégiées des mèches dans une chevelure.

– Carte directionnelle :

En combinant notre filtre fréquentiel (4.9) avec l'analyse par filtre GLOP (4.12), nous avons redéfini un nouveau filtre selon l'équation :

$$G_i(\theta) = \frac{1}{2\pi\sigma^2} \frac{1}{f} \exp\left(-\frac{(f-f_0)^2}{2\sigma^2}\right) \exp\left(-\frac{(\theta-\theta_i)^2}{2\sigma^2}\right) \quad (7.1)$$

avec σ la bande passante, f_0 la fréquence centrale et θ l'angle de l'orientation du filtre.

À partir des paramètres σ et f_0 déterminés automatiquement pour notre filtre gaussien (section 4.3), nous décomposons le spectre selon 8 orientations :

$$\theta = [0, 22.5, 45, 67.5, 90, 112.5, 135, 157.5.]$$

Nous obtenons pour chaque pixel une information de puissance de spectre associée à chaque orientation.

L'approche la plus simple est de définir la direction privilégiée comme le plan où le pixel à la plus grande puissance spectrale. Par cette méthode nous ne prenons pas en compte

les plans comprenant des puissances proches pour plusieurs directions. C'est le cas en particulier des pixels non directionnels appartenant à une zone homogène qui peut avoir la même information de puissance selon toutes les directions. En redefinissant l'intervalle de décomposition des orientations de l'intervalle $[0, \pi]$ sur l'intervalle $[0, 2\pi]$, nous obtenons ainsi une distribution homogène qui va nous permettre de définir des orientations plus précises et de calculer la somme vectorielle pour chaque pixel. Pour un pixel dans une zone homogène, l'information de direction est proche pour toutes les directions et nous pouvons retrouver un vecteur nul en effectuant la somme vectorielle (*figure 7.1*). En seuillant les vecteurs de faibles normes, nous éliminons ces pixels de notre représentation vectorielle. Pour définir la carte de champs de vecteur nous reporterons l'intervalle des directions de l'intervalle $[0, 2\pi]$ vers l'intervalle $[0, \pi]$.

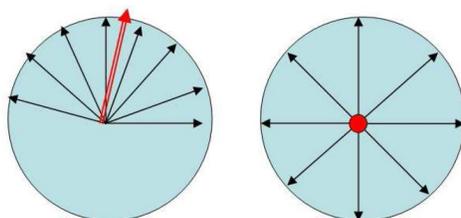


FIGURE 7.1 – *Schéma de la répartition Angulaire* : Différence de la somme vectorielle (*vecteur rouge*) des 8 vecteurs directions (*vecteurs noirs*) pour une décomposition sur l'intervalle $[0, \pi]$ (*gauche*) et sur l'intervalle $[0, 2\pi]$ (*droite*)

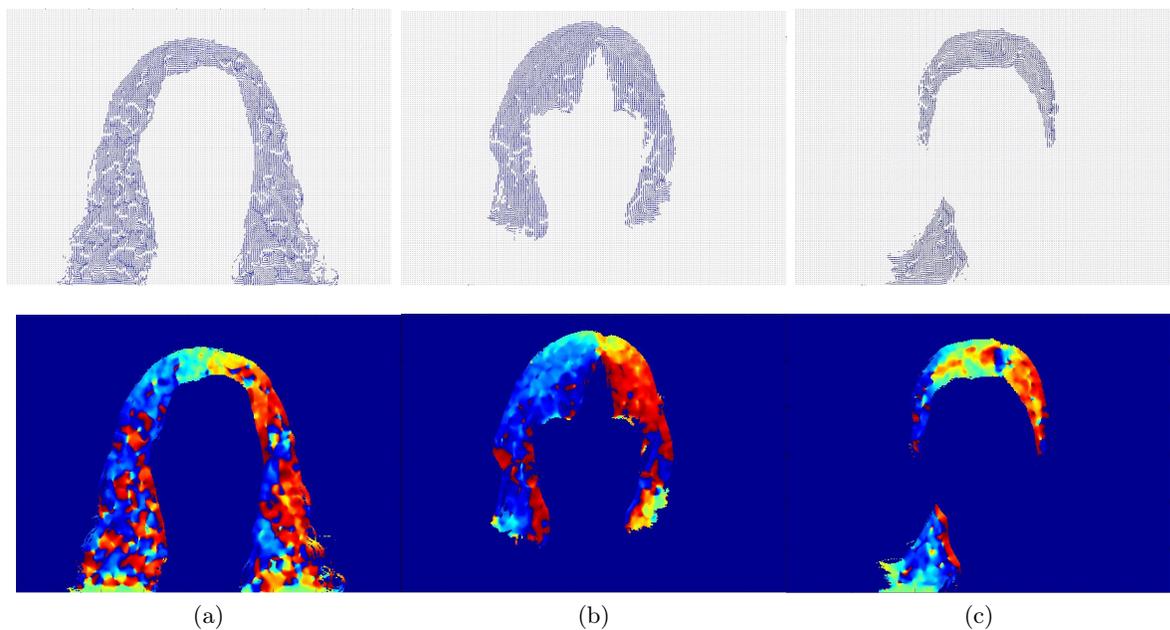


FIGURE 7.2 – **Caractérisation de la Texture - Carte directionnelle** : Carte directionnelle par une représentation vectorielle (1^{ère} ligne) et par une représentation couleur par pseudo teinte (2^{ème} ligne)

– **Caractérisation :**

Nous obtenons des cartes avec des définitions fines de la direction des cheveux pour chaque

pixel. En étudiant la répartition de la distribution du champ vectoriel, nous pouvons caractériser la forme et la répartition des cheveux et ainsi décrire certains types de coiffures. Nous pouvons définir 2 cartes :

- **Une carte du champ de vecteur de direction uniquement** : Cette carte ne contient que l'information de direction associée à chaque pixel du masque capillaire. Nous pouvons représenter cette carte de deux manières (*figure 7.2*) :
 - *Carte vectorielle* : pour chaque pixel nous modélisons un vecteur de norme constante dans la direction associée. Nous obtenons une représentation qui visuellement est similaire à une vraie chevelure avec la représentation de nombreux vecteurs que l'on peut comparer à des mèches de cheveux (*figure 7.3a*). Nous obtenons aussi une représentation fine de mèches très fines ou de cheveux fins (*figure 7.3b*).
 - *Carte couleur de pseudo teinte* : nous avons attribué une couleur à chaque composante de direction. Cette représentation est utile pour une description générale de la texture où une coloration uniforme représente une coiffure avec des cheveux lisses (*figure 7.2a*) alors qu'une représentation plus mélangée représente une coiffure bouclée (*figure 7.2b*).

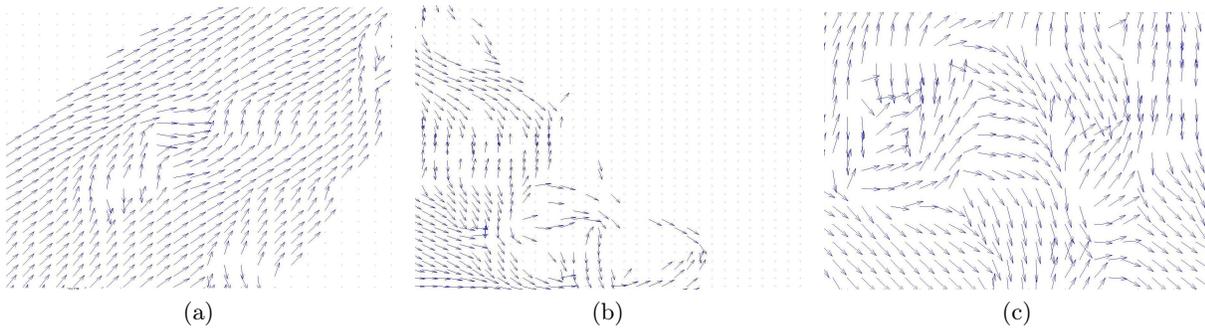


FIGURE 7.3 – **Caractérisation de la Texture - Détails de la carte de direction** : (a) Champ de direction uniforme pour une mèche ,(b) Représentation précise des mèches fines, (c) Discontinuité du champ de direction

Ces représentations donnent de bons résultats mais certaines limites à la caractérisation apparaissent. La première limite est comparable à la teinte d'une chevelure qui peut prendre plusieurs valeurs. Nous pouvons trouver plusieurs types de coiffure à l'intérieur du même masque capillaire, en particulier pour les cheveux longs où nous trouvons un type de chevelure pour le haut du masque et une autre pour le bas. La seconde limite est plus locale. Elle correspond à l'étude des changements de direction pour différencier ceux qui correspondent à une boucle de cheveux avec ceux qui représentent une mèche qui passe devant une autre (*figure 7.3c*). L'information de direction seule ne suffit pas et nous avons donc défini une nouvelle carte.

- **Une carte du champ de vecteur de direction et en puissance** : nous avons ajouté une information de puissance à l'information de direction. Elle est modélisée en modifiant la norme des vecteurs de directions. Cette carte n'améliore pas la représentation du masque capillaire, mais l'information de puissance pourra permettre la caractérisation

des zones de conflits où l'on peut trouver de nombreuses directions différentes.

Une piste possible pour améliorer la définition du champ vectoriel de la chevelure sont les travaux sur la caractérisation de texture de Germain [Germain 03] avec notamment l'utilisation de l'opérateur IRON (Isotropic and Recursive Orientation Network) qui permet une estimation des orientations locales pour des pixels qui possèdent différentes informations d'orientation pour une même échelle [Michelet 06] [Michelet 07]. Cette approche n'a pu être expérimentée faute de temps.

7.2 Caractérisation de la Teinte

Dans le chapitre 5, nous avons déterminé la position de la chevelure par l'étude de sa teinte à partir d'une analyse couleur. Nous avons défini un classifieur couleur modélisé par la distribution gaussienne du modèle couleur pour les composantes de chrominance de l'espace YCbCr.

Bien que la caractérisation de la teinte de la chevelure soit nécessaire, la définition d'un classifieur couleur n'est pas simple. Comme pour la segmentation, l'apprentissage de modèle couleur correspondant aux différents types de teinte est difficile autant pour sa définition que pour la sélection d'échantillons représentatifs pour son apprentissage.

Dans cette section, nous nous sommes surtout appliqués à définir la meilleure description couleur du masque capillaire segmenté. Nous avons mis en place un classifieur couleur simple basé sur la calcul de distance euclidienne entre les modèles couleurs de référence et le modèle couleur courant dans l'espace bidimensionnel des composantes de chrominance CbCr. Nous avons choisi 5 couleurs type pour caractériser les chevelures : Noir, brun, blond clair, blond foncé et roux.

- **Couleur moyenne du masque** : Pour chaque couleur la distribution gaussienne du modèle couleur est calculée. Dans un premier temps, seule la valeur centrale de la distribution est conservée pour déterminer sa position dans le plan CbCr. La distribution du modèle courant est calculée sur l'ensemble du masque capillaire.

Nous obtenons une caractérisation couleur qui donne de bons résultats pour des coiffures dont la distribution couleur est uniforme en particulier pour les chevelures foncées. Par contre, pour les chevelures claires plus sensibles à l'illumination nous avons des résultats plus partagés. D'une part parce que la répartition des couleurs de référence sont plus proches et d'autre part parce que l'on peut définir plusieurs teintes au sein de la même chevelure.

- **Décomposition en 3 composantes radiométriques** :

Dans leur travaux Chen *et al* [Chen 06] ont créé la carte de répartition de la lumière qui, à partir d'un filtrage passe-bas de la carte de luminance, définit les zones d'illuminations et les zones d'ombres de la chevelure. Afin de prendre cette information en compte, nous avons divisé la carte de luminance de l'image en trois sous cartes : la carte des reflets, la carte de couleur neutre et la carte d'ombres.

Pour obtenir ces cartes, nous avons utilisé une classification par un algorithme de plus proches voisins sur la distribution des pixels de luminance. Cette approche est comparable à un seuillage multimodal de l'histogramme de la carte de luminance mais ceci nous a

permis d'obtenir de meilleures cartes car nous n'avons pas toujours 3 modes séparables sur les histogrammes.

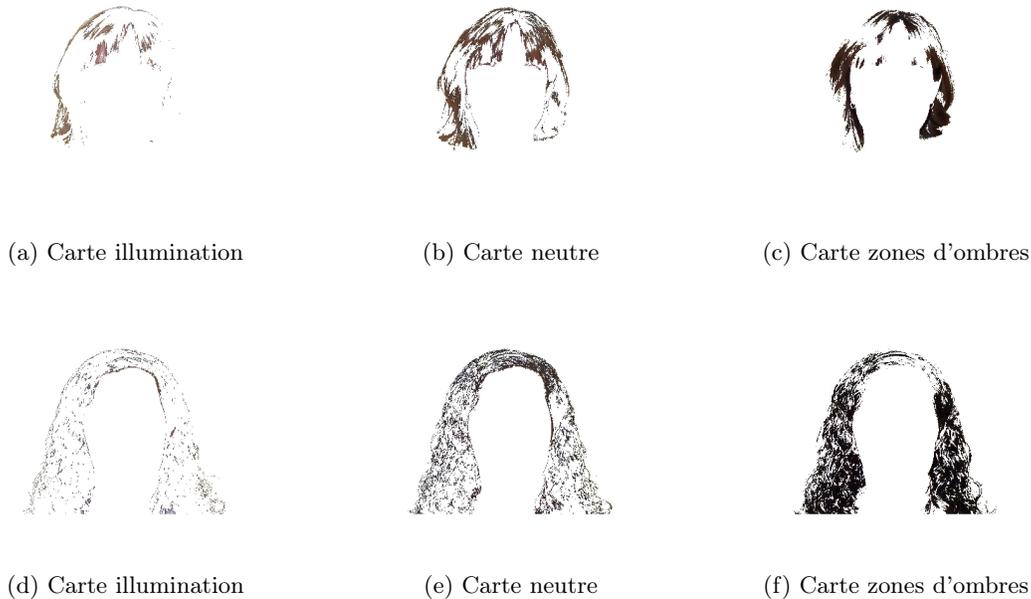


FIGURE 7.4 – **Caractérisation de la Teinte - Décomposition de la Luminance** : Cartes de répartition de la luminance pour la carte d'illumination, la carte normale et la carte des zones d'ombres pour une chevelure claire (1^{ère} ligne) et pour une chevelure foncée (2^{ème} ligne)

Nous obtenons pour chaque coiffure 3 cartes, en fonction de leur teinte nous obtenons une répartition des pixels différentes :

- *Cheveux clairs* : nous obtenons une carte qui correspond à ce que nous attendions. La carte d'illumination (*figure 7.4a*) comprend les zones de reflets spéculaires, quelques pointes de mèches qui apparaissent plus claires ainsi que certains pixels de peau. La carte neutre (*figure 7.4b*) comporte les cheveux de teintes intermédiaires. Enfin la carte des zones d'ombres (*figure 7.4c*) comprend les zones des racines et des cheveux proches du cou qui sont moins illuminés.
- *Cheveux foncés* : pour les chevelures foncées, la répartition des cartes peut être différente. Nous retrouvons bien la carte d'illuminations avec les reflets spéculaires (*figure 7.4d*). Dans notre exemple la répartition de ces zones est clairsemée car elle correspond à des reflets sur une chevelure bouclée. Nous avons aussi la carte intermédiaire (*figure 7.4e*) et la carte des zones d'ombres (*figure 7.4f*). Mais c'est dans cette dernière que nous retrouvons le maximum d'informations avec la majorité des pixels du masque capillaire.

Au final, nous appliquons le même classifieur pour chaque carte obtenue. Nous caractérisons mieux chaque carte avec un bon étiquetage dans la majorité des cas. Mais maintenant nous avons potentiellement 3 types de couleurs pour une même chevelure. Comme nous avons vu sur la répartition des cartes en fonctions du type de chevelure, le choix de la

meilleure teinte représentative n'est pas simple car ce n'est pas forcément la zone centrale qui est la plus représentative ni celle contenant le plus de pixels.

Au final le choix d'une teinte ne peut être défini comme un unique modèle mais plutôt comme la composition de plusieurs couleurs.

– **Densité de probabilité & Fonctions de croyances :**

Afin de prendre en compte la possibilité pour une chevelure d'avoir plusieurs teintes nous avons redéfini notre classifieur couleur du choix de teinte de la chevelure. Nous représentons alors la distribution gaussienne de chaque modèle couleur comme la densité de probabilité d'appartenir à ce modèle. Nous allons pouvoir définir la teinte de la chevelure comme une pondération de plusieurs couleurs. Nous aurions pu aussi utiliser le modèle de fonctions de croyances pour fusionner l'information couleur des trois cartes de la décomposition de la luminance. Nous aurions pu introduire la notion d'"incertitude" pour prendre en compte l'information couleur ou pas dans la définition de la teinte. Mais l'intégration tardive de ces fonctions dans notre algorithme de segmentation ne nous a pas permis de développer cette approche pour la caractérisation couleur.

Enfin pour optimiser cette étude, nous pouvons envisager comme pour la texture de définir une information de couleur pour chaque pixel du masque capillaire. Cette piste d'amélioration pourrait ainsi permettre une définition plus précise de la répartition de la teinte dans l'ensemble d'une chevelure et de pouvoir caractériser des types de coiffures plus complexes en particulier les coiffures avec des mèches de couleurs différentes.

7.3 Coloration du masque capillaire



(a) Coloration uniforme

FIGURE 7.5 – Coloration de la chevelure - Coloration uniforme

Une des perspectives de l'application de maquillage virtuel est de permettre le changement de coloration de la chevelure. L'utilisation de l'information récupérée sur le masque capillaire segmenté peut aider à définir des pistes de recherches pour aider à la coloration d'une chevelure. Initialement la coloration du masque capillaire nous servait seulement pour la visualisation des résultats de notre segmentation capillaire. Nous appliquons binairesment une coloration uniforme sur l'ensemble du masque capillaire segmenté. Cependant, cette représentation s'est vite avérée inadaptée car elle ne nous permettait pas de voir correctement les zones mal segmentées pour une segmentation ratée (*figure 7.5a*).

Modification de chrominance : nous travaillons principalement dans l'espace couleur YCbCr qui sépare la composante de luminance des composantes de chrominances. L'approche la plus simple pour conserver un maximum d'information de luminance est de n'appliquer le changement de teinte seulement aux composantes de chrominances. Pour une définition encore plus simple nous travaillons dans l'espace perceptif HSV (*annexe C.3*) où seul le plan de teinte H sera changé pour conserver au maximum l'information contenue par les plans de Luminosité V et de Saturation S .



FIGURE 7.6 – **Coloration de la chevelure - Coloration de chrominance** : Coloration de la chevelure pour une chevelure foncée (1^{ère} ligne) et une chevelure claire (2^{ème} ligne)

Nous obtenons de bons résultats avec des rendus plus ou moins réalistes (*figure 7.6*). Pour des coiffures claires, le changement de teinte est plus réussi car le contraste existant dans la carte de luminance est mieux conservé. Pour les chevelures foncées, les contrastes sont moins grands et l'application d'une teinte claire est moins réaliste. Nous avons tenté d'améliorer le rendu des chevelures foncées en éclaircissant les valeurs de la composante de luminosité V . En se basant sur les travaux de Chen *et al* [Chen 06] et à partir du masque d'orientation obtenu précédemment, une piste d'amélioration pour la coloration serait de combiner l'information du masque avec un traitement de synthèse afin de créer l'information de contraste pour les chevelures qui en ont peu ou pas.

Coloration par marqueurs : une autre approche qui a particulièrement retenu notre attention, est l'approche basée sur les travaux de coloration développés par Levin *et al* [Levin 04]. Initialement prévu pour la coloration de film en noir et blanc, ces travaux se basent comme pour leur travaux du traitement "*Matting*" sur la définition d'une faible nombre de marqueurs couleurs précisément placés.

Dans la figure 7.7, nous pouvons voir les bons résultats leur approche. Sur l'image de gauche,

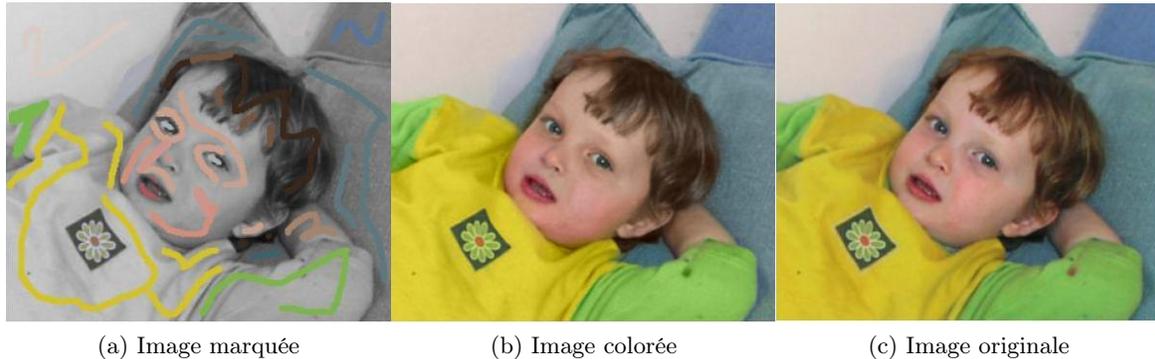


FIGURE 7.7 – **Coloration de la chevelure - Coloration par marqueurs** : Coloration similaire à l'approche "Matting" qui est initialisée par un faible nombre de pixels marqués.

nous avons l'image marquée avec les différents marqueurs couleurs, au centre l'image résultat obtenue après coloration et sur la droite, nous avons l'image couleur originale. Si nous comparons les 2 images couleurs, nous pouvons observer très peu de différence et seule une étude attentive permet de noter les artefacts de l'estimation des couleurs : une tache rouge sur le bras gauche et une légère différence de teinte de la chevelure.



FIGURE 7.8 – **Coloration de la chevelure - Coloration par marqueurs - Détails** : Détails des marqueurs pour la chevelure où on observe deux types de teintes utilisées pour colorer les cheveux

Bien que nous ayons noté une différence entre les 2 chevelures, le résultat de l'estimation couleur est très réaliste. L'autre point à noter est que l'opérateur expert qui a placé les marqueurs, a défini 2 types de marqueurs de couleurs différentes pour la chevelure (*figure 7.8*). Cet exemple montre bien le besoin de définir plusieurs teintes pour colorer une chevelure. Cette méthode donne donc de très bon résultats mais dans l'optique d'une application de coloration automatique, il est nécessaire de définir des règles qui s'appuient sur deux hypothèses :

- **Les teintes** : à partir de la coloration voulue, il faudrait définir un triplet de teintes à appliquer : la teinte principale, son éclaircissement et son assombrissement. Il faudra de plus adapter ces triplets en fonction de la coloration initiale du masque capillaire en aug-

mentant ou diminuant l'écart entre les différentes teintes.

- **Le positionnement des marqueurs** : Comme nous avons vu dans notre étude des limites du traitement "*Matting*" (section 3.3), le positionnement des marqueurs est très important pour identifier les zones à segmenter mais plus libre à l'intérieur de ces zones. Il faut donc définir des règles pour les placer au mieux automatiquement. Le positionnement des marqueurs en s'appuyant sur la décomposition en 3 composantes de luminances n'est pas suffisant, nous n'obtenons pas de bonne estimation. Pour la coloration, la position des marqueurs à l'intérieur des zones à coloriser étant importante, l'intégration de la carte de direction pourrait améliorer la définition des marqueurs.

7.4 Caractérisation de la Forme

Dans le chapitre 6, nous avons présenté notre approche de fusion de données avec l'intégration d'une fonction d'affaiblissement basée sur la position de la chevelure autour du visage. Nous avons intégré cette information par une carte isotrope autour du masque de visage. Nous admettons que la chevelure peut être présente tout autour du visage cependant sa répartition peut apporter de l'information sur le type de coupe étudiée. Yacoob *et al* [Yacoob 06] définissent 4 paramètres caractéristiques de la forme de la chevelure : sa longueur, son volume, sa symétrie et son nombre de composante.

- **La longueur des cheveux** : les cheveux peuvent être plus ou moins long allant de la coupe très courte avec une chevelure rase jusqu'aux cheveux les plus longs pouvant mesurer plus de 2 mètres. Notre descripteur doit pouvoir mesurer plusieurs longueurs et permettre la définition de l'étiquetage de longueurs types.



FIGURE 7.9 – Caractérisation de la Forme - La longueur des cheveux

- **Le nombre de composante** : généralement une chevelure se compose d'une seule composante connexe mais elle peut aussi apparaître en plusieurs subdivisions. L'origine de cette multiple représentation peut être naturelle : la perte des cheveux qui entraîne une calvitie qui donne une représentation de la chevelure en 2 parties de part et d'autre du visage. Elle peut être aussi d'origine esthétique : cette représentation est due en général pour les cheveux longs attachés en queue de cheval (2 composantes) qui apparaît sur une épaule ou pour une chevelure attachée en couettes ou nattes (3 composantes). Notre descripteur doit nous permettre de compter le nombre de composante de la chevelure et de les localiser autour du visage.



FIGURE 7.10 – Caractérisation de la Forme - Le nombre de composante

- **Le volume** : les cheveux peuvent prendre plus ou moins de volume. Cette variation peut être due à la longueur des cheveux mais aussi à leur type : une chevelure crépu prendra plus de volume qu'une chevelure lisse pour la même longueur de cheveux. Notre descripteur doit permettre de donner une mesure du volume.



FIGURE 7.11 – Caractérisation de la Forme - Le volume

Nous allons créer une signature basée sur le calcul de la répartition du masque capillaire que nous allons diviser en plusieurs zones et pour chacune d'elles compter le nombre de pixels présents. Nous obtiendrons ainsi un histogramme de répartition spatiale des pixels cheveux qui modélise l'information de volume par le nombre de pixels par zones. Nous normaliserons ce nombre par rapport à la taille du visage dans l'image. L'information de longueur est obtenue par le nombre de zones contenant des pixels. Enfin le nombre de composantes est déterminé en fonction de la répartition de ces zones.

Pour obtenir la signature, il faut donc diviser le masque capillaire segmenté. Notre première approche consiste en une division circulaire autour du visage. Nous avons calculé les coordonnées (x, y) du centre de chaque oeil grâce à un algorithme de détection des yeux. À partir de données morphologiques [Yacob 06], nous définissons le centre d'une ellipse englobant le visage selon l'équation :

$$\begin{aligned} \text{Centre Ellipse}_x &= \frac{1}{2} (\text{Oeil gauche}_x + \text{Oeil droit}_x) \\ \text{Centre Ellipse}_y &= \frac{1}{2} (\text{Oeil gauche}_y + \text{Oeil droit}_y) + 0.4(\text{Oeil droit}_x - \text{Oeil gauche}_x) \end{aligned} \quad (7.2)$$

Le centre de cette ellipse est pris comme origine de notre division circulaire. La précision de notre signature dépendra du nombre de zones créées. Il faut donc choisir un nombre suffisamment grand pour avoir assez d'information pour pouvoir classifier les chevelures. Nous avons choisi une décomposition circulaire avec un pas de 10° , ce qui nous donne une signature en 36 classes. La

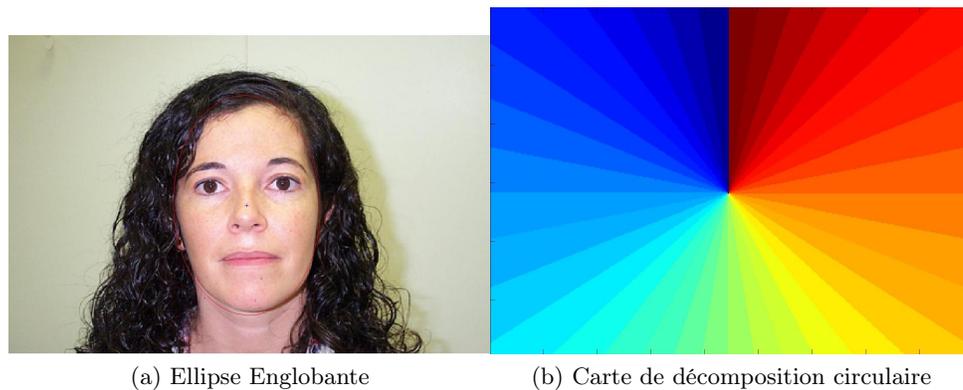


FIGURE 7.12 – **Caractérisation de la Forme - Modèle de décomposition circulaire :** Schéma de la décomposition angulaire de l'image.

décomposition s'effectue dans le sens trigonométrique avec pour origine le bas du visage pour avoir une signature symétrique par rapport au visage. En effet nous pouvons considérer l'axe médian vertical du visage comme un axe naturel de symétrie.

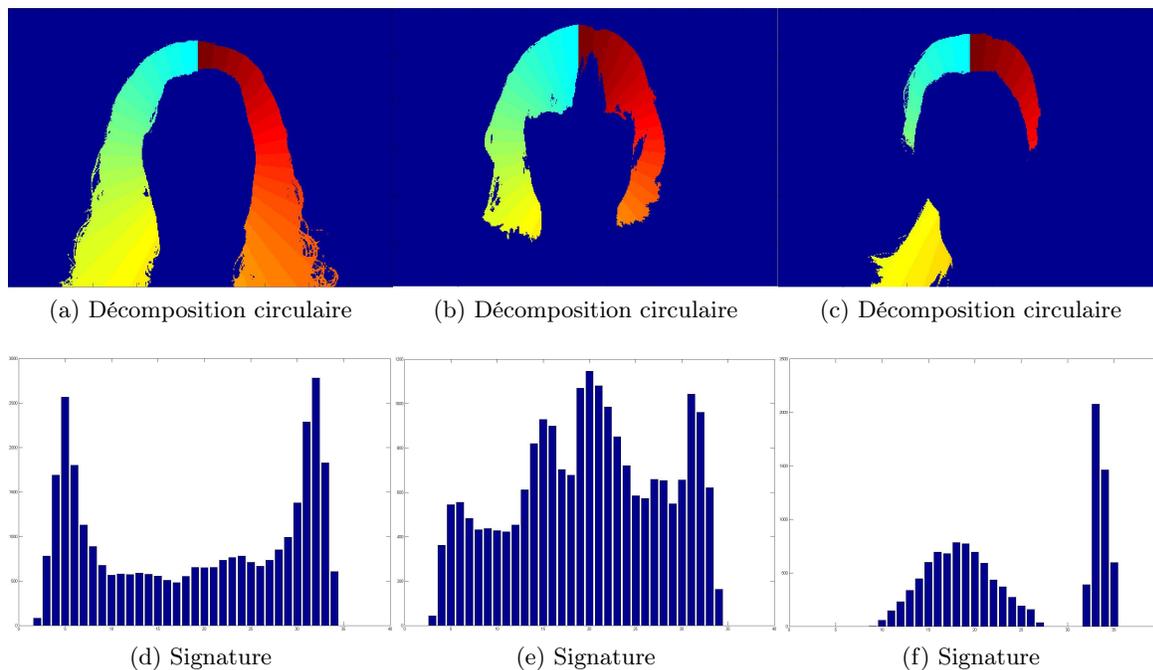


FIGURE 7.13 – **Caractérisation de la Forme - Décomposition circulaire - Résultats :** Représentation de la décomposition circulaire appliquée à des masques capillaires segmentés (1^{ère} ligne) et les signatures obtenues (2^{ème} ligne) pour 3 types de chevelure

Les premières signatures obtenues nous donnent de bons résultats et permettent de décrire la chevelure correctement. Nous pouvons déjà retrouver plusieurs détails caractéristiques de la chevelure :

- *La longueur des cheveux* : nous obtenons des signatures plus ou moins compactes avec un

certain nombre de classes ayant une valeur nulle. Plus le nombre de classes aux extrémités de la signature ont une valeur nulle plus la chevelure est courte. Nous pouvons ainsi localiser la zone de cou qui en général, quel que soit le type de coiffure, ne comporte pas de cheveux. Dans la figure 7.13d, nous obtenons une signature de 34 classes non nulle alors que nous avons une signature de 32 classes non nulle dans la signature de la figure 7.13e. Nous avons donc une chevelure plus longue dans la première figure. Résultat que l'on peut aussi déduire de la répartition de classes avec une grande valeur aux extrémités non nulles de la signature.

- *Le volume* : nous pouvons observer la distribution de pixels par zones de la décomposition. Dans la figure 7.13d nous observons bien un volume constant pour les classes du sommet du crâne avec une augmentation des valeurs pour les cheveux longs. Pour la figure 7.13d, nous avons une répartition plus homogène sur l'ensemble du masque capillaire. On peut observer aussi un trou dans la distribution pour le sommet du crâne qui correspond bien à la coupure dans la zone de la frange.
- *La symétrie* : Nous pouvons aussi définir une information de symétrie selon la répartition de la distribution de la signature. Chaque extrémité de la signature représente les pixels autour du cou. Dans la figure 7.13d, nous avons une chevelure bien symétrique. Pour la figure 7.13e nous avons une coiffure comportant des dissymétries : une sur le sommet du crâne et une aux niveaux des extrémités que nous pouvons expliquer respectivement par la position de la frange et par la répartition des cheveux longs. Enfin dans la figure 7.13f nous observons une dissymétrie principalement sur la répartition des différentes composantes qui permet de placer le côté de la queue de cheval par rapport au cou.
- *Le nombre de composante* : nous observons bien la composition de la chevelure en une ou plusieurs composantes connexes. Nous définissons chaque zone de chevelure entre chaque classe de valeur nulle de la signature. Nous observons bien des coiffures à une seule composante connexe pour les figures 7.13d et 7.13e. Pour la figure 7.13f, nous avons bien une coiffure en deux parties. Par la localisation avec un bloc au sommet du crâne et un bloc sur le côté, nous pouvons aussi bien caractériser le type de coiffure comme une chevelure avec une queue de cheval et non pas une calvitie où les deux blocs auraient été placés symétriquement.

Même si nous obtenons de bons résultats, nous voyons immédiatement une limite à la définition circulaire de la division du masque capillaire. Pour les cheveux longs, nous avons deux pics autour de la zone de cou qui devrait traduire une augmentation du volume de la chevelure. Hors ce n'est pas forcément le cas, cette augmentation est surtout due à la représentation de chaque zone comme portion angulaire de l'image. Plus on s'éloigne du centre de la décomposition, plus nous obtenons une grande zone de projection (figure 7.14).

Pour contrer cet effet nous avons redéfini la carte de division du masque capillaire. Nous nous sommes basé sur une représentation inspirée de l'implantation capillaire. Nous conservons la distribution circulaire pour la partie supérieure du masque capillaire. Pour la zone inférieure, nous séparons la zone en bandes parallèles séparées au centre par la droite passant par le centre du visage. Dans cette zone, nous trouvons généralement les cheveux long qui "*tombent*" le long du visage de part et d'autre du cou. Cette nouvelle représentation va nous permettre de mieux décrire les cheveux longs. Nous conservons notre représentation en 36 classes pour avoir autant

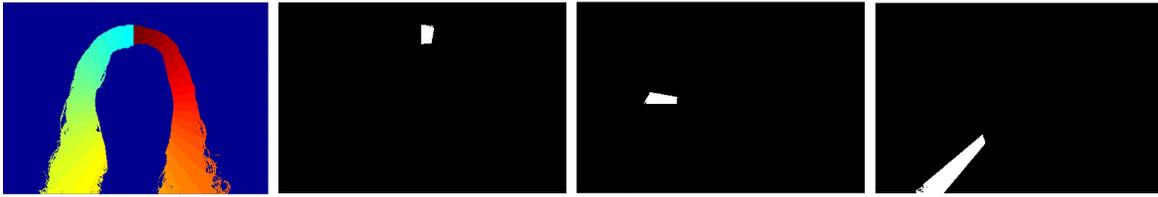


FIGURE 7.14 – **Caractérisation de la Forme - Projection Circulaire** : Représentation de la projection angulaire pour différentes orientations possibles.

de classes pour la partie supérieure que la partie inférieure de la chevelure. Nous définissons la largeur de nos bandes par rapport à la largeur de la taille du visage. En fonction de la taille du visage et de son placement dans l'image, nous pouvons obtenir des signatures qui n'ont pas 36 classes. Pour un visage de petite taille nous pouvons avoir une signature plus longue avec plus d'information pour les zones du bas du cou et du haut du torse, nous tronquons la signature pour conserver une signature en 36 classes. Pour un visage de grande taille, nous avons des signatures plus courtes avec moins de 36 classes, nous les complétons avec des classes à valeur nulle.

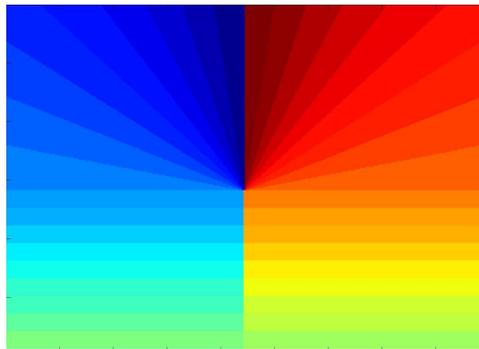


FIGURE 7.15 – **Caractérisation de la Forme - Modèle de décomposition semi-circulaire** : Représentation du nouveau modèle de la décomposition du masque capillaire inspirée par l'implantation des cheveux sur le visage. Le modèle est circulaire pour la partie supérieure du masque capillaire et en bandes parallèles pour la zone inférieure

Nous obtenons des signatures semblables à celle obtenue pour la décomposition circulaire. Ces résultats s'expliquent en partie car la partie supérieure de la nouvelle décomposition est aussi circulaire. Pour la partie inférieure, nous obtenons une meilleure caractérisation particulièrement des cheveux longs avec une meilleure répartition du volume sur l'ensemble des classes :

- *La longueur des cheveux* : la différence entre les cheveux courts et cheveux long est plus précise et nous pouvons vraiment donner une mesure de comparaison de la différence de longueur des cheveux. Dans la figure 7.16d, l'ensemble des classes représentent les cheveux, les classes des extrémités de la signature ne représentent plus la zone de cou. L'amélioration de la caractérisation de la longueur est surtout visible sur la signature de la queue de cheval dans la figure 7.16f.
- *Le volume* : la répartition du volume est aussi mieux caractérisée avec le retrait des pics que l'on trouvait pour la décomposition capillaire. Dans la figure 7.16d la répartition est

plus homogène sur l'ensemble du masque et traduit mieux l'augmentation du volume pour les pointes. Ces résultats sont aussi visibles pour la représentation de la queue de cheval 7.16f dont la grosseur grandit bien avec la distance, information qui n'était pas correcte sur la décomposition circulaire.

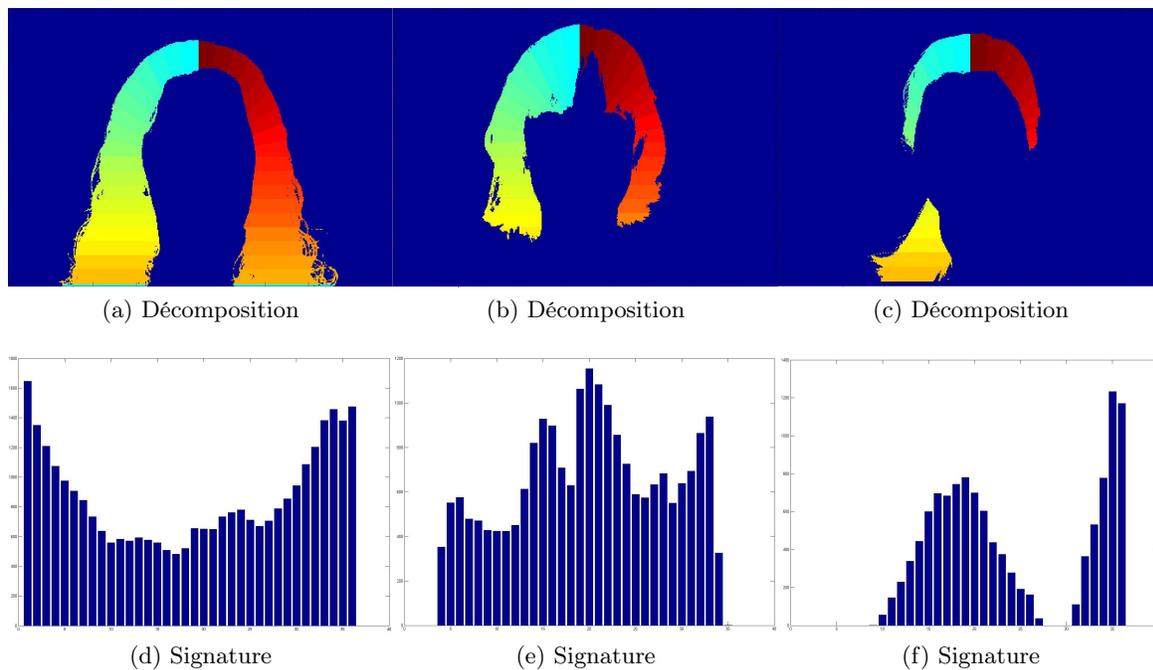


FIGURE 7.16 – **Caractérisation de la Forme - Décomposition semi-circulaire - Résultats** : Représentation de la décomposition semi-circulaire appliquée à des masques capillaires segmentés (1^{ère} ligne) et les signatures obtenues (2^{ème} ligne) pour 3 types de chevelure

Nous avons défini une signature pour chaque chevelure qui décrit bien sa répartition autour du visage. L'étape suivante de la caractérisation est la définition des modèles types de signature pour la classification. Bien que nous n'ayons pas l'expertise pour définir les classes pour une application biométrique, nous avons essayé de voir si l'information contenue dans les signatures permettait de définir des classes. Nous avons effectué une analyse par composante principale (ACP) sur les signatures obtenues pour 160 images de la base Caltech afin de voir si l'information portée par la signature était assez discriminante. Nous obtenons plus de 70% de l'information portée par la première composante mais nous obtenons 95% de l'information pour les 13 premières composantes. Nous avons bien une description qui contient de l'information. La question qui se pose est de savoir si cette information peut permettre une classification utilisable pour la classification de la chevelure.

Dans nos exemples, nous avons présenté trois signatures bien distinctes permettant de définir facilement trois classes. Pour un grand nombre de coiffures, cette classification est plus difficile à mettre en place et nous trouvons beaucoup de signatures intermédiaires. Nous avons été confrontés à plusieurs cas où nous obtenons des signatures différentes bien que, après une observation subjective, nous nous attendions à avoir des chevelures similaires. Ces différences se retrouvent principalement selon deux paramètres :

- *La différence de longueur* : nous observons deux cas de figure. Pour le premier cas, nous avons des signatures de la même longueur mais pour des implantations différentes. Inversement pour le deuxième cas, nous avons des longueurs différentes pour la même implantation. Ainsi une information de texture pourrait permettre une amélioration de la comparaison des signatures.
- *La variation de volume* : nous calculons l'information de volume à partir du nombre de pixels présents dans chaque zone. Cette information ne permet pas d'observer la différence de volume entre une frange et un chignon qui vont entraîner un grand volume sur la partie supérieure de la chevelure. Une information de position à l'intérieur ou l'extérieur de l'ellipse englobante du visage pourrait aussi permettre une amélioration de la comparaison des signatures.

7.5 Conclusion

Dans ce chapitre, nous avons introduit une première étude sur la caractérisation de la chevelure. À partir des problèmes rencontrés pour la caractérisation et la localisation des 3 paramètres caractéristiques de la chevelure, nous avons proposé des pistes de recherches pour la définition de descripteurs.

Nous avons caractérisé la texture par la définition du champ de vecteur décrivant l'orientation des cheveux dans la chevelure. Pour la caractérisation de la teinte, nous nous sommes plus intéressés à la manière de définir la teinte de la chevelure plutôt que de créer un classifieur optimal. Nous avons essayé de prendre en compte la possibilité pour une chevelure d'avoir plusieurs teintes. Enfin nous avons associé une signature de la répartition des pixels cheveux autour du visage pour caractériser sa longueur, son volume, sa symétrie et son nombre de composante.

Toutes ces pistes doivent être étudiées plus précisément. Nous avons déjà présenté plusieurs perspectives de recherches futures notamment l'intégration des fonctions de croyances pour la définition de la teinte et l'utilisation de la coloration par marqueurs. La caractérisation capillaire à elle seule pourrait faire l'objet de futurs travaux de thèse tant le travail de recherche restant à effectuer est encore important. Il faudrait rajouter en parallèle une composante biométrique et cognitive beaucoup plus importante que celle que nous avons pu mettre en place dans nos travaux.

Conclusion

Dans ce manuscrit de thèse nous avons présenté l'ensemble de nos travaux sur la segmentation du masque capillaire dans un visage. La chevelure est une nouvelle composante des travaux menés au GIPSA-lab qui peut apporter de réelles solutions pour un système de maquillage virtuel ou des applications de caractérisation ou de modélisation d'un visage. Les objectifs visés étaient de présenter une approche pour la segmentation automatique du masque capillaire dans une image. À partir de l'étude de la représentation des paramètres caractéristiques de la chevelure, nous devons modéliser la position de la chevelure avec le plus de robustesse et de précision possible. Par robustesse, nous entendons une approche automatique pouvant traiter le plus grand nombre de coupes de cheveux possibles. Par précision, nous entendons la définition fidèle d'un masque de position de la chevelure.

Dans cette conclusion, nous proposons de dresser un bilan de notre étude sur la segmentation capillaire. Puis nous proposerons plusieurs perspectives de recherches pour poursuivre nos travaux, résoudre certains problèmes soulevés par notre étude et explorer certaines pistes que nous n'avons pas eu le temps de traiter au cours de cette thèse.

Bilan de la segmentation capillaire

Pour commencer nous avons étudié l'importance de la chevelure dans la représentation d'une personne : elle est le reflet de sa personnalité et une composante principale de son identité. L'étude de la littérature traitant du sujet nous a permis de définir 3 paramètres caractéristiques que nous avons employés dans notre algorithme : la texture, la couleur et la localisation de la chevelure. Nous avons aussi fait une revue des méthodes de segmentation, les techniques dites par approches "*contours*" et les techniques dites par approches "*régions*". Cette étude nous a permis de mettre en place une segmentation par traitement "*Matting*", segmentation par diffusion de marqueurs dont la position est initialisée grâce aux 3 paramètres de la chevelure. Ce traitement permet la segmentation d'une image en 2 plans, un plan d'information \mathbf{F} qui représente pour nous le masque capillaire et un plan de fond \mathbf{B} . Ces plans sont liés par un plan nommé alpha matte qui est estimé durant le traitement.

Au travers de ce rapport, nous avons présenté notre méthode de segmentation capillaire que nous illustrons par un organigramme récapitulatif de la boucle de notre algorithme (*figure 7.17*).

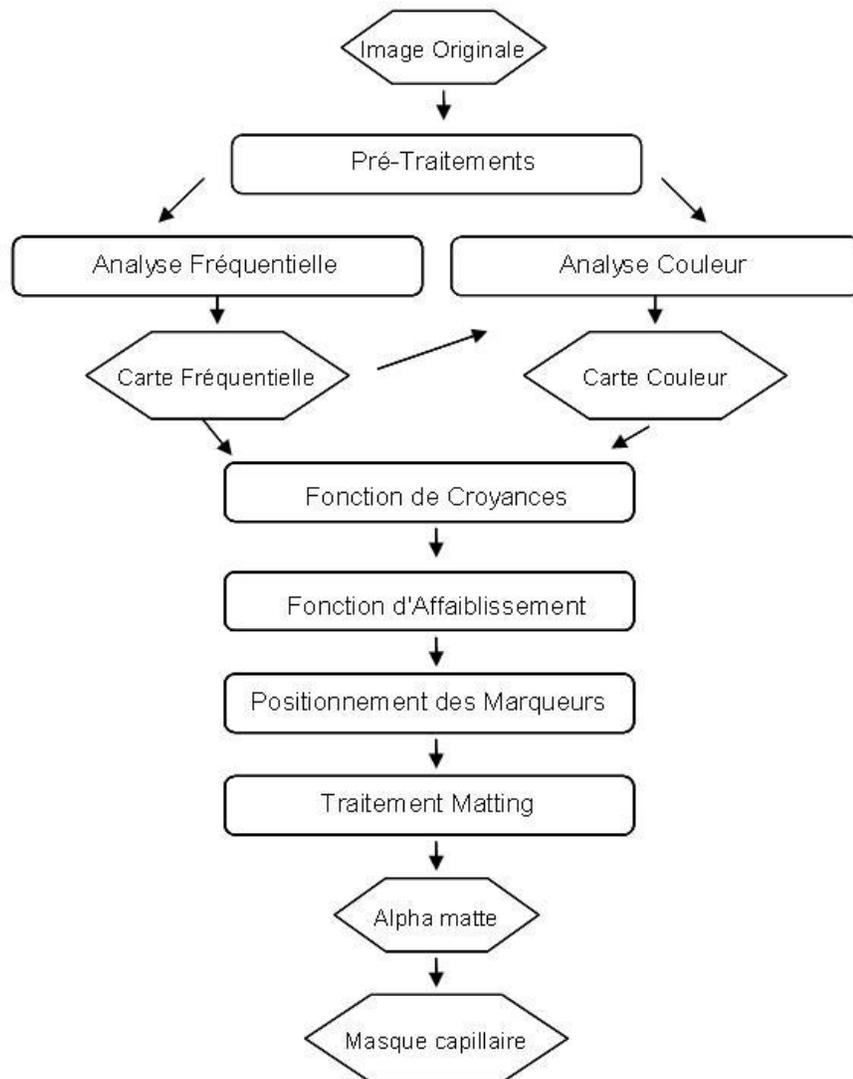


FIGURE 7.17 – Organigramme de la segmentation du masque capillaire dans un visage

Phase d'initialisation

Cette phase comprend tous les pré-traitements nécessaires pour préparer l'image à la segmentation capillaire. Elle est initialisée par la recherche du visage du modèle à étudier par l'algorithme de Viola & Jones combiné avec la détection de l'emplacement des yeux. Cette localisation du visage permet l'étude couleur des pixels du visage afin de segmenter le masque de couleur de peau. Pour améliorer les contrastes de l'image un redressement de la luminance aura été effectué par un traitement basé sur la représentation numérique d'un modèle de rétine.

Une fois l'image pré-traitée, nous pouvons commencer l'étape de segmentation du masque capillaire proprement dite.

L'analyse fréquentielle

Dans cette partie, nous avons analysé l'image en fréquence afin de localiser les zones contenant de l'information fréquentielle similaire à celle de la chevelure. Dans un premier, nous nous sommes inspirés de la modélisation de la rétine par une décomposition en fréquence et en orientation de l'image. Comme l'information de direction n'est pas utile avant la caractérisation capillaire, nous avons redéfini notre analyse et nous avons ainsi mis en place un traitement fréquentiel par un filtre multi-directionnel modélisé par une couronne gaussienne. Nous avons aussi étudié la texture particulière de la chevelure par une analyse fractale. Enfin nous avons aussi présenté nos travaux pour l'automatisation de la définition des paramètres de notre filtre : la fréquence centrale et la bande passante.

L'analyse couleur

Dans cette partie, nous avons analysé l'image afin de localiser les zones contenant de l'information couleur similaire à celle de la chevelure. Nous avons, dans un premier temps, comparé les différents espaces couleurs afin de choisir l'espace le plus stable possible pour les grandes variations de teinte d'une chevelure. Puis dans un second temps, nous avons étudié différents classifieurs couleurs afin de choisir celui dont la représentation de la distribution couleur était la plus compacte. Nous avons ainsi mis en place une segmentation couleur par un classifieur couleur dans l'espace YCbCr avec une distribution couleur modélisée par une gaussienne pour chaque composante de chrominance.

La fusion de données

Dans cette partie, nous avons présenté notre approche pour combiner l'information de localisation de la chevelure en fréquence et couleur. Suite à une étude préliminaire par fusion logique, nous avons mis en place une fusion de données par le modèle des fonctions de croyances transférables. Ce modèle permet une définition plus fine des marqueurs du traitement "*Mattting*" et notamment de contrôler le placement des pixels dont l'état est inconnu et qui reste à être défini par l'estimation de l'alpha matte. Afin de prendre en compte l'information de "haut niveau" de la position de la chevelure par rapport au visage, une fonction d'affaiblissement a été ajoutée à notre modèle de fusion. Cette fonction permet la pondération de la croyance qui est diminuée en fonction de la distance au visage. Plusieurs implémentations ont été mises en place afin de contrôler les zones de sur-segmentation de l'image.

La caractérisation de la chevelure

Dans cette partie, nous avons analysé le masque capillaire segmenté pour le caractériser selon les 3 paramètres de la chevelure. La caractérisation de la texture est effectuée par le calcul d'un champ de vecteur représentant la direction des cheveux. La caractérisation couleur quant à elle consiste à l'étiquetage en fonction des couleurs type d'une chevelure. Enfin la caractérisation de la localisation s'appuie sur la définition d'une signature basée sur la répartition des pixels cheveux autour du visage.

Perspectives

Ces travaux de thèse ouvrent de nombreuses perspectives, des améliorations pour notre approche de segmentation, des pistes d'extensions de recherches pour la caractérisation du masque capillaire et l'intégration du logiciel de maquillage virtuel.

Segmentation capillaire

À court terme, nous pouvons proposer plusieurs pistes qui permettraient d'améliorer la méthode de segmentation proposée :

- *l'optimisation de la définition automatique des paramètres de notre filtre fréquentiel* : nous avons vu que la définition de ces paramètres étaient difficile en ne se basant que sur l'étude fréquentielle de l'image. Afin d'améliorer leur définition la prise en compte des autres paramètres de la chevelure pourrait permettre de définir des règles d'arrêt pour la boucle itérative.
- *l'amélioration de la définition de la carte de distance* : nous avons vu que nous pouvions contrôler la segmentation du masque capillaire en fonction de la définition de la carte de distance dans notre cas avec la prise en compte ou non de la zone du cou. Une caractérisation des résultats pourrait permettre l'automatisation du choix de la carte à utiliser pour obtenir le meilleur masque capillaire et limiter la sur-segmentation de l'image.
- *Une seconde boucle d'optimisation du masque capillaire* : nous avons orienté notre algorithme pour généralement privilégier une sur-segmentation du masque capillaire plutôt qu'une sous-segmentation. La mise en place d'une seconde boucle de traitement pour affiner la définition de la chevelure segmentée. L'utilisation des descripteurs pourra guider cette approche afin que le masque capillaire obtenu tende vers un modèle de chevelure répertorié.

À long terme, il serait envisageable d'assouplir l'hypothèse sur la pose du visage. Nous nous sommes limités tout au long de notre étude à des images où le visage était vu de face, pose qui n'est pas toujours facile à contrôler en fonction des images traitées (en particulier pour la base mannequins dans notre cas). De plus la caractérisation de la rotation du visage pourra permettre aussi l'amélioration de la définition de la carte de distance qui pourra être anisotrope. Enfin la prise en compte de la pilosité faciale est une autre piste à étudier.

Caractérisation capillaire

À court terme, la définition de catégories pour classifier les paramètres de texture et de localisation du masque capillaire segmenté pourrait permettre de décrire la coupe selon les 3

descripteurs que nous avons définis dans la section 2.1.7. À long terme, le travail combiné avec un expert en cognition pourra permettre une meilleure définition de la représentation de ces paramètres de description. Une étude sur la répartition des cheveux par rapport à un visage permettra aussi de déterminer des règles sur l'implantation capillaire pour l'adaptation automatique d'un masque de synthèse à un visage afin d'obtenir des rendus réalistes

Application au maquillage virtuel

À court terme, la première perspective est l'implémentation de l'algorithme en langage C pour l'intégrer au logiciel Makeuponline de la société Vesalis. Cette extension à l'application de maquillage virtuel pourra permettre, dans un premier temps, la coloration de la chevelure. À plus long terme, l'implémentation de la synthèse de coiffure est prévue. Une étape d'optimisation du code pourra rendre le traitement plus rapide et une meilleure gestion de la mémoire pourra permettre d'augmenter la résolution des images traitées. Une autre piste de développement est d'utiliser les paramètres de caractérisation de la chevelure pour adapter le maquillage au mieux aux modèles traités.

Annexe

A Le Cheveu



Sommaire

A.1	Le Cheveu	181
A.2	Structure	182
A.2.1	La Racine	182
A.2.2	La Tige	183
A.2.3	La Kératine	184
A.3	Cycle de vie	185
A.4	Type de Chevelure	186
A.5	Coloration	187
A.5.1	Coloration naturelle	187
A.5.2	Canitie	189
A.6	Propriétés du Cheveu	190
A.6.1	Élasticité et Solidité	190
A.6.2	Effet Éponge	190
A.6.3	Électricité	190
A.6.4	Propriété Optique	191

Nous avons étudié la chevelure et sa représentation dans une image. Dans cette annexe nous nous intéressons aux propriétés structurelles et physiques d'un cheveu. Pour cela nous nous inspirerons largement des travaux de L'Oréal Hair science [Science] et des travaux de thèse de Florence Bertails [Bertails 06].

A.1 Le Cheveu

La comparaison entre le poil et le cheveu est commune et cette analogie est renforcée par la définition lexicale du mot cheveux. Les définitions que l'on trouve sur les différents dictionnaires en ligne définissent : **cheveu** *N.m.* du latin *capillus* Poil qui pousse sur le crâne des êtres humains.

De nombreux invertébrés (en particulier les insectes) et vertébrés possèdent des poils dont les fonctions sont très diverses : pelage à poils creux protecteur, à capteur de chaleur et de rayonnement UV chez l'ours polaire (*figure A.1a*) aux poils de moustaches du chat (*figure A.1b*) ou vibrisses de la chausse souris qui jouent un rôle dans l'équilibre et la perception. Presque tous les insectes sont aussi pourvus de poils microscopiques dont on ne connaît pas exactement les fonctions. Certaines espèces d'araignées ou certaines chenilles (*figure A.1c*) sont très velues et peuvent posséder une pilosité urticante ou allergène.

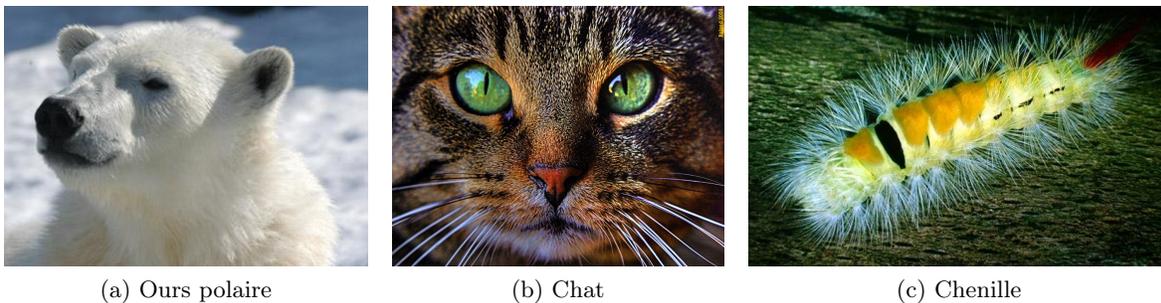


FIGURE A.1 – Les poils dans le règne animal

Les vestiges et réminiscences de notre ancien pelage sont les réflexes instinctifs. Les petits muscles lisses appelés muscles arrecteurs ou horripilateurs permettent aux poils de se dresser face aux stimuli reçus : En général la peur où chez les animaux ce réflexe sert à impressionner l'ennemie en paraissant plus gros ou le froid avec la réaction dite de "chair de poule".

Si le poil est commun à beaucoup d'espèce, le cheveu est bien le propre de l'homme. En effet contrairement au poil, le cheveu ne comporte pas ou peu de moelle responsable de la thermorégulation de l'organisme de l'animal. Même si les raisons et origines exactes de la chevelure demeurent encore un mystère de nos jours, certaines hypothèses l'expliquent soit par la nécessité de protection du crâne soit par une évolution résultant de la stature verticale de l'homme. Une chevelure humaine est composée, en moyenne, de 100 000 à 150 000 cheveux implantés seuls ou par groupes de deux ou trois dans le cuir chevelu. Le nombre total de cheveux sur une tête dépend de plusieurs facteurs : le diamètre et la couleur du cheveu, l'origine ethnique, l'âge de la personne et la prédisposition génétique à perdre ou non ses cheveux prématurément. La densité capillaire (en moyenne 200 à 300 cheveux par centimètre carré) est donc variable d'une personne

à l'autre, mais aussi d'une région du cuir chevelu à l'autre chez un même individu (en général, la densité capillaire est plus forte sur le dessus du crâne que près des tempes).

A.2 Structure

Le cheveu est composé de deux parties principales : la tige, seule partie visible (et morte) du cheveu et la racine, partie vivante encastrée profondément dans le cuir chevelu.

A.2.1 La Racine

La racine, appelée follicule pileux, est constituée de 6 parties : la matrice pileuse, la papille dermique, les kératocytes, les mélanocytes, la gaine épithéliale interne et la gaine épithéliale externe. Le follicule pileux est le siège d'une intense activité biochimique et métabolique qui conduit au développement et à la croissance de la tige pileuse. En effet, contrairement aux végétaux, la pousse du cheveu s'effectue par sa base.

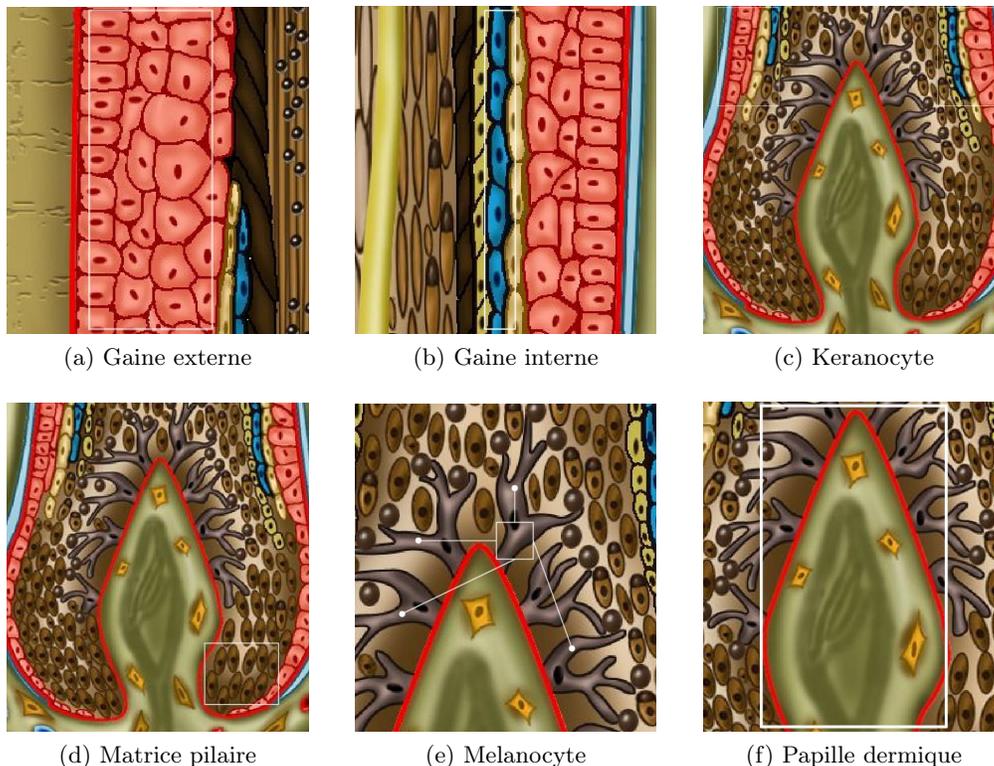


FIGURE A.2 – Constitution de la racine d'un cheveu

- **La matrice pileuse** (*figure A.2d*) : zone la plus profonde du composant épithélial du follicule. C'est ici que selon les instructions de la papille dermique, les kératocytes se multiplient et commencent à se différencier pour former les différents compartiments du follicule : les gaines épithéliales internes et externes et la tige pileuse. Dans le même temps, les kératocytes se chargent de kératine et ceux qui constitueront la moelle et le cortex de la fibre capillaire reçoivent des pigments colorés transmis par les mélanocytes.

- **La papille dermique** (*figure A.2f*) : résultat de l'association et l'interaction d'un composant épithélial et d'un composant dermique. Enfouie à quatre millimètres dans le cuir chevelu, elle est formée de cellules conjonctives qui sécrètent une importante matrice extracellulaire, très vascularisée, c'est le véritable moteur du follicule. C'est elle qui fournit les informations nécessaires à la multiplication et à la différenciation des cellules de la matrice pileuse régissant ainsi les cycles de vie du cheveu.
- **Les Kéranocytes** (*figure A.2c*) : sont les cellules constitutives aussi bien que l'usine biochimique constituée par le follicule que du produit fini qui en résulte : la tige pileuse. Pour ce faire, ils se multiplient au sein de la matrice à un rythme beaucoup plus élevé que celui des kéranocytes de la peau. D'autre part, ils se différencient pour former différentes structures du follicule que sont les gaines épithéliales internes et externes et la tige pileuse. Dans le même temps, ces cellules se chargent de kératine qu'elles fabriquent elle-même. Cette kératinisation provoque leur durcissement et aboutit à la désintégration de leur noyau et à leur mort.
- **Les Mélanocytes** (*figure A.2e*) : fabriquent les pigments de mélanine et les transportent aux kéranocytes qui formeront le cortex de la tige pileuse. Le cheveu est ainsi coloré dès son origine. Ce sont de grosses cellules arrimées au sommet de la papille dermique par un pôle basal. Contrairement à ceux de la peau, les mélanocytes du follicule n'ont pas besoin de la lumière solaire pour produire de la mélanine. D'autre part, alors que ceux de la peau fournissent une quarantaine de kéranocytes, ceux du cheveu n'en servent que quatre ou cinq.
- **La gaine épithéliale interne** (*figure A.2b*) : entoure la tige pileuse en formation. À ce niveau, la kératine est encore amorphe et non organisée. La tige est molle et fragile, la gaine interne lui sert de tuteur. Elle joue aussi un rôle de « barrière », vis-à-vis des messages venant de la papille dermique. Ceux-ci doivent être « triés » afin de préserver les gaines épithéliales internes et externes et la tige pileuse.
- **La gaine épithéliale externe** (*figure A.2a*) : En continuité de l'épiderme, elle entoure la gaine interne. Elle abrite des cellules pluripotentes activées par la papille dermique lors de la reprise du cycle de croissance du cheveu et permet alors la reconstruction du follicule. Elle contient aussi dans sa partie supérieure un « réservoir » de mélanocytes réactivés eux aussi lors de la reprise de la croissance du cheveu.

A.2.2 La Tige

La tige est constituée de trois parties : la cuticule, le cortex et la moelle.

- **La moelle** (*figure A.3a*) : ou canal médullaire est situé au centre de la fibre. Son diamètre varie entre dix et vingt micromètres. Les cellules qui la composent dégénèrent rapidement pour laisser place à une bulle d'air. Son rôle précis est encore inconnu chez l'homme. Chez l'animal sa structure alvéolée assure le rôle essentiel de la régulation thermique.
- **La cuticule** (*figure A.3b*) : enveloppe externe de la fibre capillaire. Les cellules qui la composent sont appelées écailles. Incolores, d'une longueur de cinquante à soixante-dix micromètres pour une largeur de cinq à dix micromètres et ne dépassant pas le micro-

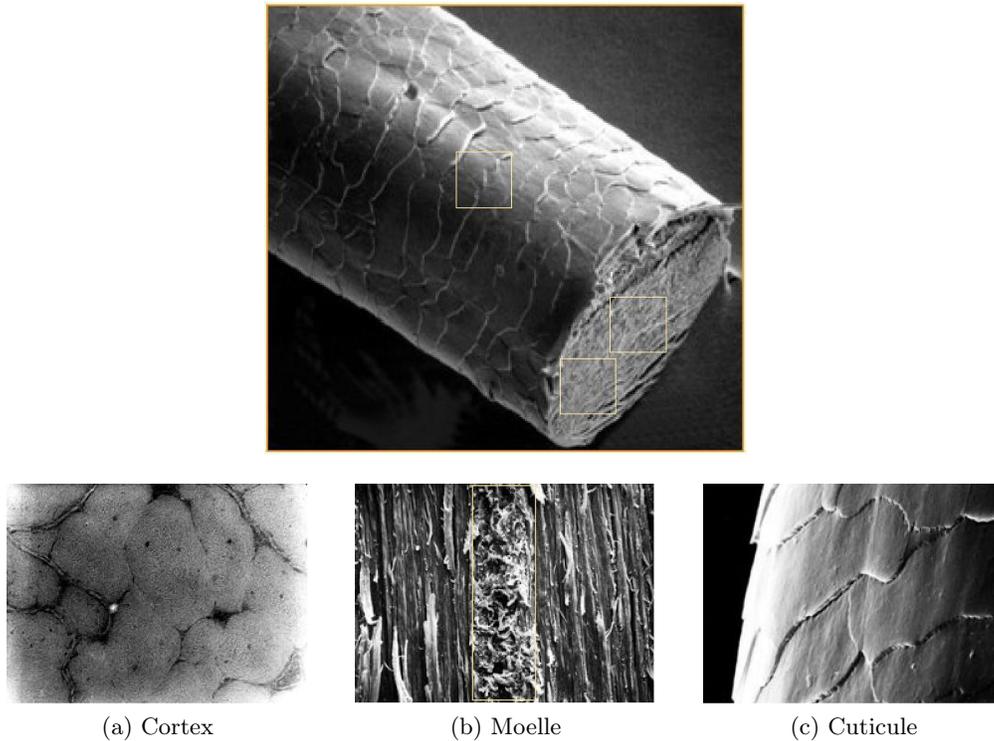


FIGURE A.3 – Constitution de la tige capillaire

mètre d'épaisseur. Reliées par un ciment cellulaire riche en lipides, elles se superposent les unes aux autres à la manière d'une toiture en ardoise. Finalement la cuticule est formée de trois à dix épaisseurs. Avec ses écailles dont le bord libre est orienté vers l'extrémité du cheveu, elle peut ainsi assurer sa fonction de protection du cortex. C'est en effet elle qui subit les agressions du cheveu et c'est de son état que dépend la santé de la fibre capillaire.

- **Le Cortex** (*figure A.3c*) : est le corps de la fibre capillaire. Il représente 90 % du poids total d'un cheveu. Il est formé de cellules remplies de kératine dont l'organisation confère à la fibre des propriétés remarquables. Au cours de leur trajet depuis la matrice, ces cellules corticales se sont allongées et atteignent une centaine de micromètre pour un diamètre de trois à six micromètres. Disposées dans le sens de la longueur du cheveu, elles sont incluses dans un ciment intercellulaire composé d'une kératine souple et riche en lipides. C'est également dans le cortex que l'on trouve les grains de mélanine qui donnent sa couleur au cheveu. Ses seules protections naturelles sont une fine couche de sébum et la cuticule. Un Cheveu est constitué d'eau de lipides, de traces d'éléments minéraux de mélanine et de Kératine qui représentent près de 95 % du cheveu.

A.2.3 La Kératine

La Kératine est une protéine constituée de 28 acides aminés synthétisées et utilisées par un grand nombre d'être vivant comme élément de structure. C'est aussi la protéine fibreuse type. Il existe deux types de kératine : L'alpha-kératine présente chez les mammifères dont l'homme et la bêta-kératine présente chez les reptiles et les oiseaux. On peut encore distinguer deux

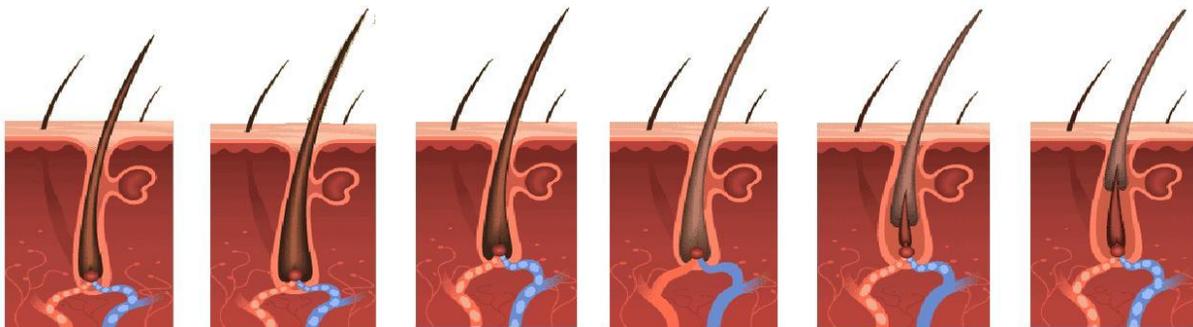
types d'alpha-kératines. Les kératines épithéliales ou cytokératines qui constituent la peau et les kératines exokératines qui constituent les poils et les cheveux.

A.3 Cycle de vie

Chaque cheveu suit un cycle de vie appelé cycle pileux. Ce cycle est composé de quatre phases : la phase anagène, la phase catagène, la phase télogène et la phase de latence. Au cours de notre vie chaque cheveu va suivre en moyenne une vingtaine de cycles complets. C'est un phénomène unique dans le corps humain, le processus pileux se re fabrique à l'instar du processus de déroulant dans l'embryon.

Afin d'avoir une chevelure constante, les cheveux suivent leur cycle de vie asynchrones et indépendamment des autres cheveux. Chaque cheveu est absolument autonome. Le fait de conserver une chevelure de densité constante s'explique par la proportion de cheveux en phase de croissance (environ 90 %). Néanmoins, même si les cheveux en phase de chute ne constituent que 10 % de l'ensemble de la chevelure, il est courant de perdre une centaine de cheveux par jour.

- **Phase anagène** (en moyenne 3 ans) : phase de croissance du cheveu. Elle dure en moyenne 3 ans mais peut varier de quelques mois à 10 ans selon les individus. Le bulbe pileux établit une communication avec la papille dermique. Les messages échangés engendrent la création de kératinocytes dans le fond du bulbe. La pousse d'un cheveu est le résultat visuel de la multiplication cellulaire. En effet c'est la pression exercée par les kératinocytes qui pousse le cheveu vers l'épiderme et le fait apparaître. Contrairement aux végétaux, le cheveu pousse par la racine. Il possède le taux de multiplication cellulaire le plus important du corps humain et pousse d'environ 0.3 mm par jour soit en moyenne 1 cm par mois. Pour une chevelure normale, cela revient à une production d'environ 15 km de cheveux en un an.
- **Phase catagène** (3 semaines 1 mois) : phase de repos. Les kératinocytes du bulbe dégèrent et meurent. Le bulbe remonte vers la surface de l'épiderme en se séparant de la papille dermique. Cette séparation coupe les échanges et entraîne la "mort" du cheveu.
- **Phase télogène** (2 à 3 mois) : Le cheveu remonte dans l'épiderme où il va tomber. Cette chute est accélérée par les frottements de toutes sortes que constituent entre autres le brossage et le lavage.
- **Phase latence** (2 à 5 mois) : Aucune activité n'est visible. Puis suite a un stimulus dont l'origine demeure encore un mystère, un nouveau cycle reprend et une nouvelle phase d'anagène commence.



A.4 Type de Chevelure

Un cheveu peut être de différentes formes : raide, ondulé, frisé ou crépu. Mais sa composition de base est toujours la même. C'est sa forme dans l'espace qui varie énormément. Elle dépend en grande partie de la forme de sa section et de la façon dont il pousse. Ces deux éléments semblent de plus en plus être indissociables de la forme du follicule pileux et de son implantation dans le cuir chevelu. On peut dire que cette "usine biologique" fonctionne comme un moule, en donnant sa forme au cheveu dès sa fabrication. La fibre capillaire étant une matière biologiquement morte, elle conserve ensuite cette forme originelle.

La section d'un cheveu n'est jamais tout à fait circulaire, c'est généralement une ellipse qui peut tendre vers le cercle ou nettement s'aplatir. Une analogie avec d'autres matériaux montre comment la forme de cette section influe sur l'apparence du cheveu dans l'espace : en effet, à condition d'avoir des tailles du même ordre de grandeur, un ruban presque plat s'enroulera beaucoup plus facilement sur lui-même qu'une ficelle cylindrique.

On peut définir trois grands types de cheveux associés à trois types d'ethnies : type Africain, Caucasic et Asiatique. Ce classement est bien sûr simplifié et ne reflète pas la diversité capillaire existante. Seulement même si on peut différencier une chevelure de type écossaise d'une chevelure de type suédoise ou sicilienne, les différences restent minimes.

- **Le cheveu de type asiatique** : est le cheveu qui pousse le plus rapidement avec en moyenne 1,3 cm par mois. Inversement c'est la chevelure possédant la densité la plus faible. Les cheveux poussent droit et perpendiculairement au cuir chevelu. Les cheveux asiatiques sont raides et presque ronds et de surcroît assez gros.



(a)

FIGURE A.4 – Le cheveu de type asiatique

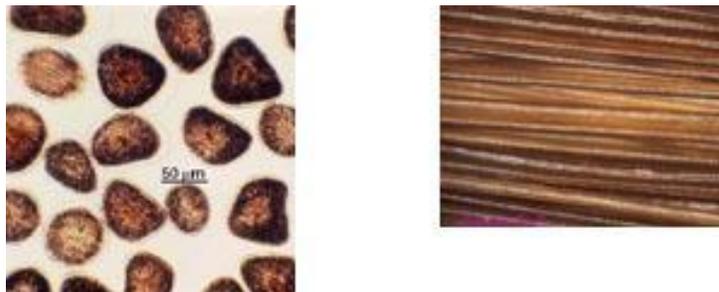
- **Le cheveu de type africain** : Avec une croissance de moins de 0,9 mm par mois, c'est le cheveu qui pousse le plus lentement. Il possède une densité légèrement supérieure au cheveu de type asiatique. Les cheveux poussent presque parallèlement au cuir chevelu en se vrillant sur lui-même. Les cheveux africains, présentant une section assez aplatie et plus fine, formeront des chevelures crépues, dont les boucles peuvent former des cercles très serrés, d'un diamètre de quelques millimètres seulement.



(a)

FIGURE A.5 – Le cheveu de type africain

- **Le cheveu de type caucasien** : est le cheveu avec une croissance moyenne avec 1,2 cm par mois. Par contre c'est la chevelure possédant la plus forte densité. Les cheveux poussent de façon oblique et présentent une certaine courbure. Les cheveux caucasiens, quant à eux, admettent une plus grande variabilité et, leur section étant une ellipse plus ou moins accentuée, les chevelures caucasiennes pourront être raides ou fortement bouclées en passant par l'ondulées.



(a)

FIGURE A.6 – Le cheveu de type caucasien

Les scientifiques ont remarqué qu'il existait une forte corrélation entre la frisure, le diamètre et l'ellipticité d'un cheveu : les gros cheveux (de type japonais par exemple) ne sont jamais bouclés à l'état naturel, les cheveux bouclés ont quant-à-eux, généralement, une section elliptique. Par contre, les cheveux les plus frisés ne sont pas nécessairement ceux qui ont la section la plus elliptique. Mais un cheveu, quel qu'il soit, n'est jamais de taille constante sur toute sa longueur. À cet égard, le cheveu asiatique, nettement plus homogène dans sa longueur, se présente comme le plus régulier et le cheveu africain comme le plus irrégulier.

A.5 Coloration

A.5.1 Coloration naturelle

La mélanine est produite par les 'mélanocytes'. Ce sont de grandes cellules en forme d'étoiles dont les branches sont appelées les dendrites. Ils produisent des mélanosomes (granules pigmentées riches en mélanines). Ensuite, en allongeant leurs dendrites, elles injectent ces pigments aux kératinocytes de la tige pileaire en formation. Ainsi, dès son origine, le cheveu acquiert sa couleur,

une couleur faite pour durer tout le temps de sa vie. Cela est d'autant plus extraordinaire quand on sait qu'un cheveu ne contient pas plus de 1 % de mélanine !

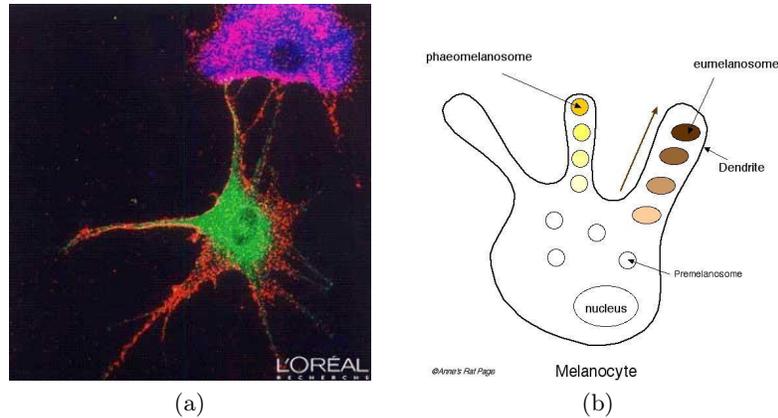


FIGURE A.7 – La Mélanine

On connaît deux principaux types de mélanine comportant chacun deux sous-groupes :

- *L'eumélanine* : se présente sous la forme d'un petit grain de riz dont la couleur va du brun-rouge au noir foncé. Elle résulte de la transformation d'un acide aminé essentiel, la tyrosine, par une enzyme, la tyrosinase. On peut distinguer deux couleurs : le type noir qui est la forme la plus oxydée de la mélanine et le type brun qui est issu du type noir mais ayant subi une mutation durant son oxydation.
- *La phaéomélanine* : possède une forme moins précise et se présente sous forme de tâche diffuse. Sa couleur va du jaune au rouge. Si elle est différente de l'eumélanine, c'est que, outre la tyrosine, sa fabrication fait intervenir un autre acide aminé, la cystéine, riche en soufre.

Le dosage de ces deux mélanines détermine la couleur du cheveu. Mais, s'il est aisé de comprendre que le cheveu noir japonais ne contient quasiment que de l'eumélanine alors que le cheveu roux irlandais est très riche en phaéomélanine. Il est plus surprenant de découvrir que le cheveu blond scandinave est constitué lui aussi principalement d'eumélanine. Cela est lié à l'immensité de déclinaisons possibles dans le mélange entre les deux pigments, déclinaison aussi bien en termes de qualité que de quantité. Ainsi la répartition des mélanines, déterminée par le patrimoine génétique de chacun, offre une palette infinie allant du blond le plus clair au noir le plus profond.

Se colorer les cheveux est devenu un acte ordinaire avec facilité d'utilisation des produits de coloration des cheveux. Nous pouvons distinguer 3 types de colorations :

- **La coloration permanente** : elle ne s'élimine pas par les shampooings et son action dure jusqu'à la repousse de la racine des cheveux. La coloration permanente permet de colorer en plus clair, dans le même ton, en plus foncé ou d'apporter des reflets. Elle est également capable de colorer une chevelure entièrement blanchie.
- **La coloration temporaire** : elle s'estompe en douceur au fil des shampooings. Elle s'utilise pour colorer dans le ton naturel ou en un peu plus foncé, pour apporter des reflets

ou pour colorer une chevelure qui présente jusqu'à 30% de cheveux blancs bien répartis.

- **La coloration fugace** : elle permet d'obtenir, jusqu'au shampoing suivant, une légère modification de la couleur naturelle et d'y apporter des reflets

A.5.2 Canitie

Le blanchissement des cheveux s'appelle la canitie. C'est un phénomène très complexe dont la première explication est très simple : dès qu'un cheveu n'est plus pigmenté lors de sa conception, il pousse blanc.

On a longtemps cru que cela était la conséquence naturelle de l'arrêt de production de mélanine par les mélanocytes. Et puis les chercheurs ont fait une découverte étonnante au sein du bulbe pileux du cheveu blanc : s'il on y trouve en effet des mélanocytes incapables de produire des pigments colorés, il en existe d'autres en parfait état de fonctionnement mais ne parvenant plus à transmettre leur mélanine aux kératinocytes. Aujourd'hui, les raisons de cette interruption de communication entre mélanocytes et kératinocytes sont encore obscures.

Une autre découverte a ensuite été faite : alors qu'on pensait que les mélanocytes étaient localisés uniquement au fond du bulbe pileux, on s'est rendu compte qu'il en existait un réservoir situé plus haut dans la gaine épithéliale externe. Ces mélanocytes sont endormis : ils ne produisent pas de pigments. Ce sont certains d'entre eux que le follicule pileux recrute pour repeupler sa partie profonde lorsqu'il commence à se reconstituer à la fin de la phase télogène. Une fois sélectionnés, ces mélanocytes sont réactivés et la production de mélanine reprend. Or, ce réservoir existe encore dans le follicule des cheveux blancs. Cela conduit à penser que la canitie pouvait être liée à un "déficit de recrutement" dans ce réservoir. Là encore, les processus mis en jeu sont mystérieux.

On peut rapprocher la canitie de l'albinisme qui est une anomalie génétique dont l'affection se caractérise par une absence de pigmentation de la peau, des poils, des cheveux, des yeux rouges, due à l'absence de mélanine.

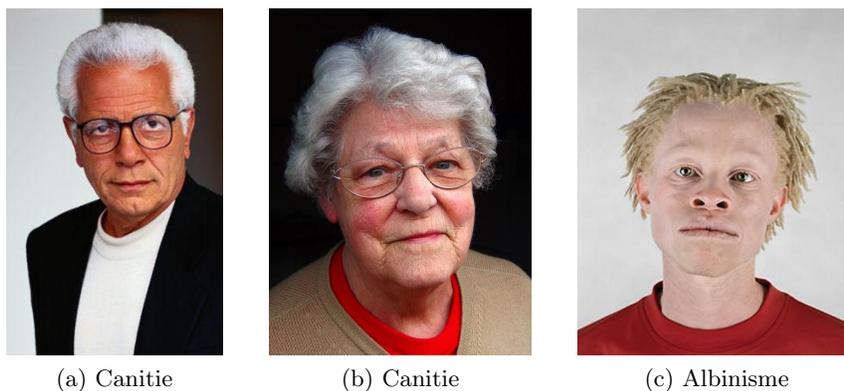


FIGURE A.8 – Canitie et Albinisme

A.6 Propriétés du Cheveu

A.6.1 Élasticité et Solidité

Si cela n'est peut-être pas flagrant quand on manipule un seul cheveu, il suffit d'essayer d'en rompre une petite mèche pour être convaincu : le cheveu est extrêmement solide. L'organisation de la kératine au sein de son cortex lui permet de supporter une traction pouvant aller jusqu'à une centaine de grammes. Une mèche de 100 cheveux résistera donc à un poids de 10 kilogrammes. Quant à une chevelure moyenne, elle pourrait soutenir 12 tonnes, si le cuir chevelu y résistait lui aussi !

Avant de rompre, le cheveu subit des transformations. Par exemple, la manipulation délicate d'un cheveu assez long montrera aisément qu'il se comporte d'abord comme un élastique. Après une faible extension, il retrouve sa longueur d'origine comme le ferait un ressort. Ainsi, jusqu'à 5% d'allongement, le cheveu est élastique. Ensuite, le cheveu entre dans une zone dite d'écoulement où, presque sans effort, il peut s'allonger de 25%. Au-delà, dans une phase précédant la rupture, le cheveu peut encore s'allonger et il casse souvent alors que sa longueur a déjà doublé. En cassant sous une traction de 60 grammes après un allongement de 40%, le cheveu africain apparaît comme le plus fragile. À l'autre extrémité, le cheveu asiatique est le plus solide, supportant un poids de 100 grammes et un allongement de 55%. Que ce soit pour l'un ou l'autre de ces aspects, le cheveu caucasien occupe une place intermédiaire.

Passée la phase d'élasticité, le cheveu présente une autre propriété : il conserve, pour un moment du moins, la forme qu'on lui donne. Ainsi, si l'on enroule un cheveu autour d'un stylo et qu'après quelques heures on enlève le stylo, le cheveu gardera sa forme bouclée. C'est ce qu'on appelle la plasticité du cheveu. Combinée à l'eau et la chaleur, cette propriété permet de modifier provisoirement la forme du cheveu.

A.6.2 Effet Éponge

Les cheveux mouillés sont plus lourds que les cheveux secs : cela paraît évident et ça l'est. Cependant, cette constatation simple illustre une caractéristique importante du cheveu : il est perméable. Malgré les écailles serrées de sa cuticule et le sébum qui l'enduit naturellement, un cheveu en bon état est susceptible d'absorber plus de 30% de son propre poids en eau. Si le cheveu est déjà abîmé par d'autres facteurs, ce pourcentage peut atteindre 45%. Sa longueur peut alors augmenter de 2% et son diamètre de 15% à 20%.

A.6.3 Électricité

Le contact des cheveux avec certains vêtements synthétiques peut aller jusqu'à produire des étincelles. Plus couramment, il suffit parfois d'enlever un pull-over pour voir nos cheveux se dresser sur notre tête. De la même manière, une règle en plastique frottée sur un tissu puis amenée à proximité des cheveux aura tendance à les attirer. Ces phénomènes sont liés à la capacité qu'a le cheveu de se charger en électricité statique. La kératine étant un excellent isolant, ce sont principalement les frottements qui électrisent ainsi le cheveu : cet effet est dit triboélectrique. Il augmente lorsque la cuticule du cheveu est abîmée, favorisant alors l'échange de charges électriques. Il diminue par contre quand le cheveu est humide : la règle en plastique approchée de cheveux mouillés ne produira aucun effet. En revanche, certaines chevelures seront très difficiles à coiffer par temps sec.

A.6.4 Propriété Optique

La composition et la structure microscopique des cheveux jouent un rôle prédominant vis-à-vis de leur apparence visuelle. La fibre capillaire peut être considérée comme un milieu translucide, d'indice de réfraction constant égal à 1,55 [Robbins 02]. Comme nous l'avons vu précédemment, c'est le cortex de la fibre capillaire qui contient les grains de mélanine, responsables de la couleur des cheveux. Ces pigments absorbent la lumière réfractée à travers le cheveu et donc plus un cheveu est pigmenté, plus la quantité de lumière absorbée est importante.

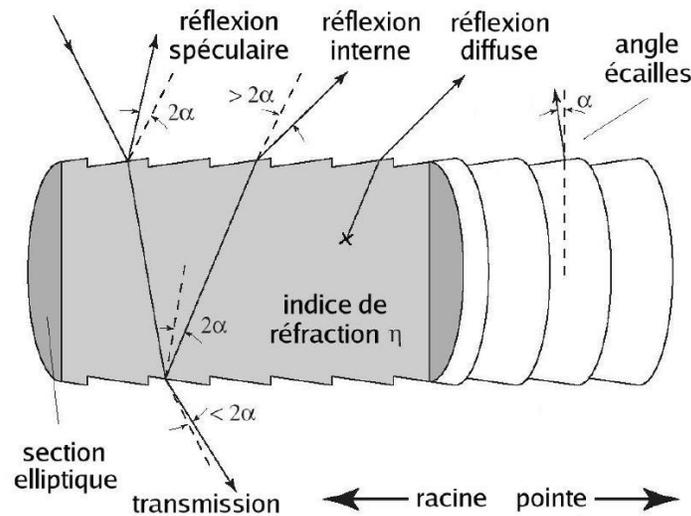
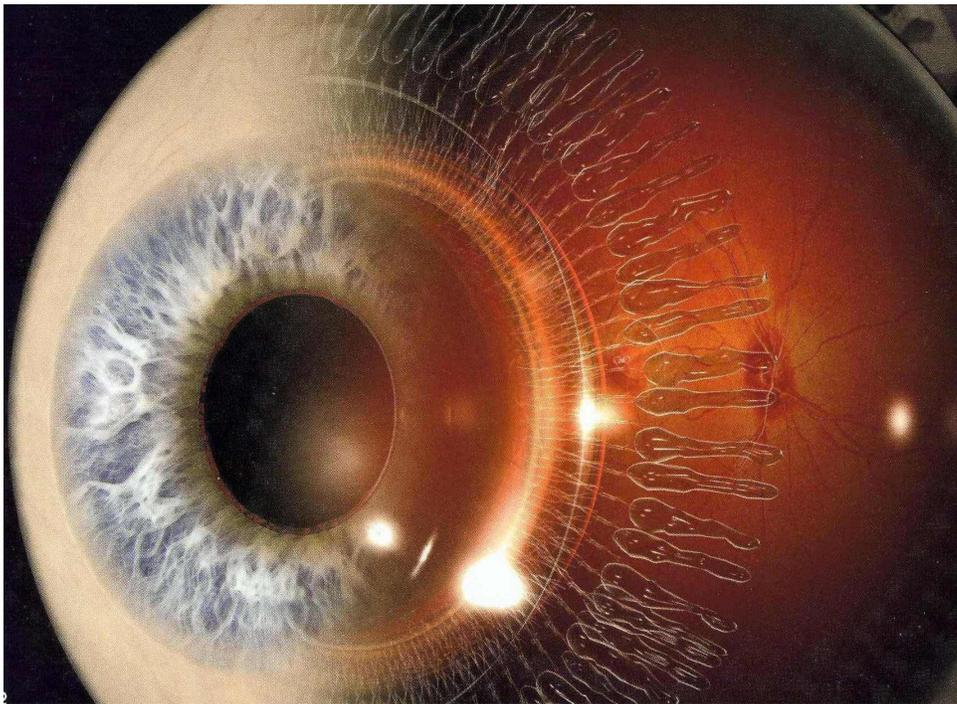


FIGURE A.9 – Modèle optique d'un cheveu[Marschner 03]

La structure particulière de la surface du cheveu, composée d'écailles qui se recouvrent partiellement, a une influence très importante sur la manière dont le cheveu interagit avec la lumière [Marschner 03]. Comme pour un cylindre semi-transparent, une portion de la lumière se réfléchit directement sur la surface de la fibre (réflexion spéculaire, de la couleur de la lumière), tandis qu'une autre pénètre par réfraction à l'intérieur de la fibre où elle est absorbée, puis restituée en partie vers l'extérieur (réflexion diffuse, colorée par les pigments de la fibre). Mais à la différence du cylindre parfait, la présence des écailles sur la surface de la fibre dévie légèrement les rayons lumineux réfléchis (phénomène de dispersion).

B La Rétine



Sommaire

B.1 Les photo-récepteurs	196
B.2 La Couche Plexiforme Externe	196
B.3 La Couche Plexiforme Interne	197
B.4 Cortex <i>V1</i>	198

La rétine est l'organe sensible de la vision. Elle fait partie de l'oeil qui modélise le système d'acquisition du Système Visuel Humain (SVH). L'iris et le cristallin représentent le système optique. L'iris dont le diamètre est variable permet le contrôle de la quantité de lumière entrant dans l'oeil tandis que le cristallin permet l'ajustement de la focale de l'oeil.

La rétine est une mince membrane d'environ 0,5 mm d'épaisseur couvrant environ 75% de la face interne du globe oculaire. Elle permet de convertir l'information lumineuse en signaux nerveux électriques. L'image est alors analysée au sein même de la rétine en zones de contraste et le résultat est envoyé au reste du système visuel par le nerf optique. Dans la suite de cette section, nous présenterons la rétine selon les travaux de Hérault [Hérault 01] et Benoit [Benoit 07].

Elle est constituée de 3 couches de neurones (*figure B.1*) :

- **La couche des photo-récepteurs** : Bâtonnets et Cônes.
- **La couche granuleuse interne** : Cellules horizontales, Cellules bipolaires et Cellules amacrine.
- **La couche ganglionnaire** : Cellules ganglionnaires.

On pourra remarquer que la répartition des cellules au sein de la rétine est inversée à leur ordre de traitement

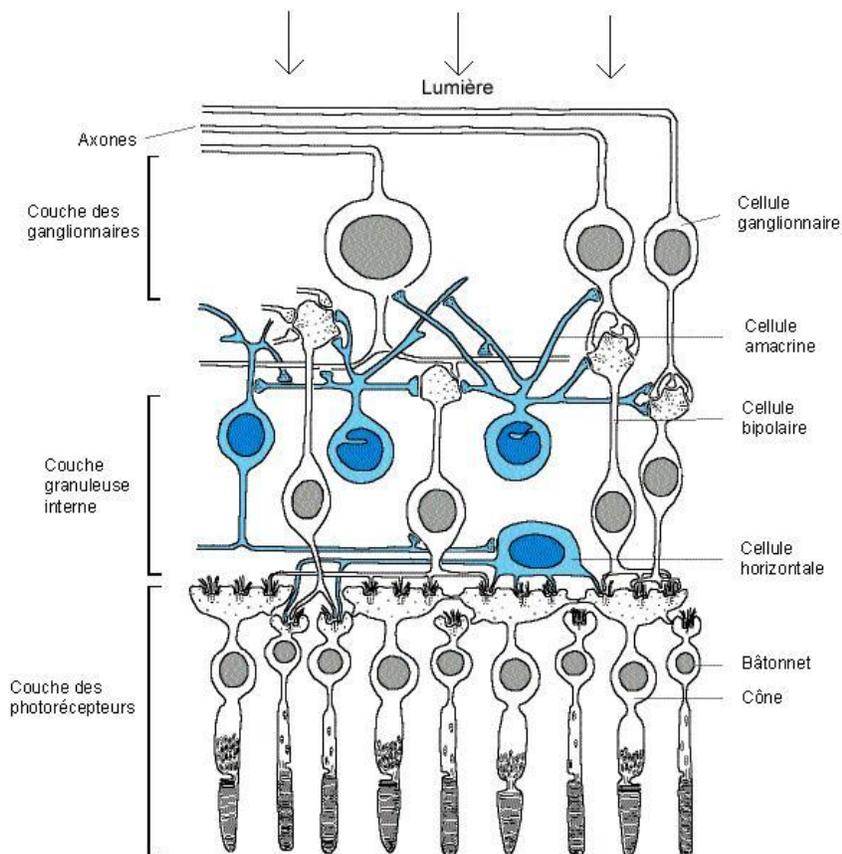


FIGURE B.1 – **Schema de l'organisation spatiale de la rétine** : Représentation des 3 couches de neurones constituant la rétine

B.1 Les photo-récepteurs

Ils constituent la couche profonde de la rétine. Leur rôle est de coder l'excitation lumineuse sous forme d'un potentiel de membrane. Ils sont répartis en 2 types : les Bâtonnets qui représentent 95% des photo-récepteurs. Ils sont sensibles aux faibles intensités lumineuses et sont dédiés à la vision nocturne (scotopique). Les cônes représentent seulement 5% des photo-récepteurs et sont dédiés à la vision diurne (photopique). Les cônes et les bâtonnets ne sont pas répartis de façon uniforme sur la rétine. Les cônes sont fortement concentrés autour de la fovéa (zone centrale de la rétine).

Les photo-récepteurs sont connectés aux cellules bipolaires et aux cellules horizontales ainsi qu'entre eux, ce qui a pour effet de lisser l'information visuelle. En effet ils sont capables d'adapter leur réponse à la luminance moyenne ambiante. Cette capacité est appelée compression adaptative logarithmique. Leur sensibilité moyenne se translate et se centre sur la luminance locale moyenne, leur dynamique restant la même. Ainsi le comportement général de la dynamique d'un photo-récepteur peut être représenté par une loi logarithmique.

B.2 La Couche Plexiforme Externe

La couche plexiforme externe **PLE** correspond à la zone de jonction entre les photo-récepteurs, les cellules horizontales et les cellules bipolaires.

- *Les cellules horizontales* : elles sont en contact entre les photo-récepteurs et les cellules bipolaires. Elles relient plusieurs photo-récepteurs et sont connectées entre elle par l'intermédiaire de synapse. L'information transmise sera lissée par le grand nombre de connexions entre une seule cellule horizontale et de nombreux photo-récepteurs d'une zone de la rétine. Ainsi chaque cellule horizontale transmettra une information sur la luminance moyenne $L(p)$ qui permettra par rétroaction d'effectuer la compression adaptative des photo-récepteurs.
- *Les cellules bipolaires* : elles relient les photo-récepteurs et les cellules horizontales aux cellules ganglionnaires. Il existe 2 types de cellules bipolaires les cellules bipolaires **ON** et les cellules bipolaires **OFF**. On peut diviser le champ récepteur de la rétine en 2 zones : une zone centrale où les cellules bipolaires **ON** seront plus sensibles et une zones périphérique où les cellules bipolaires **OFF** sont plus sensibles. Les cellules bipolaires étant reliées aux photo-récepteurs et aux cellules horizontales, on pourra définir la réponse des photo-récepteurs (respectivement cellule horizontales) comme le signal du champ récepteur central (respectivement périphérique). Les cellules bipolaires **ON** et **OFF** travaillent de manière complémentaire , en effet les cellules ne pouvant coder que des signaux positifs il faut pouvoir conserver les parties positives et négatives du signal visuel. Ainsi la réponse des cellules bipolaires **ON** (respectivement **OFF**) correspondent à la différence entre les sorties des photo-récepteurs (respectivement cellules horizontales) et des cellules horizontales (respectivement photo-récepteurs) à laquelle on a annulé les zones négatives.

La modélisation de la rétine est une représentation au cours du temps. Pour travailler sur des images statiques où l'effet de la temporalité n'a pas d'effet, nous étudierons les effets de la **PLE** pour une fréquence temporelle nulle.

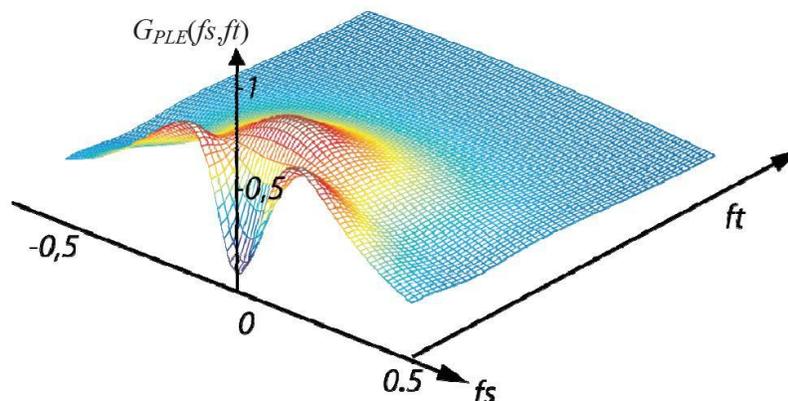


FIGURE B.2 – Fonction de transfert de la PLE [Bouvier 10]

La **PLE** a alors un effet passe bande (*figure B.2*). Le réseau des photo-récepteurs réalise un premier filtrage passe bas spatial avec une fréquence de coupure haute élevée afin d'atténuer le bruit entre les photo-récepteurs. Le réseau des cellules horizontales effectue un 2^e filtrage passe bas avec une fréquence de coupure plus basse que celle des photo-récepteurs. On obtient alors une estimation de luminance locale moyenne. En combinant les informations des sorties des cellules bipolaires **ON** et **OFF**, on obtient une image sur laquelle la luminance moyenne locale et le bruit de haute fréquence ont été supprimés et sur laquelle les contours ont été renforcés.

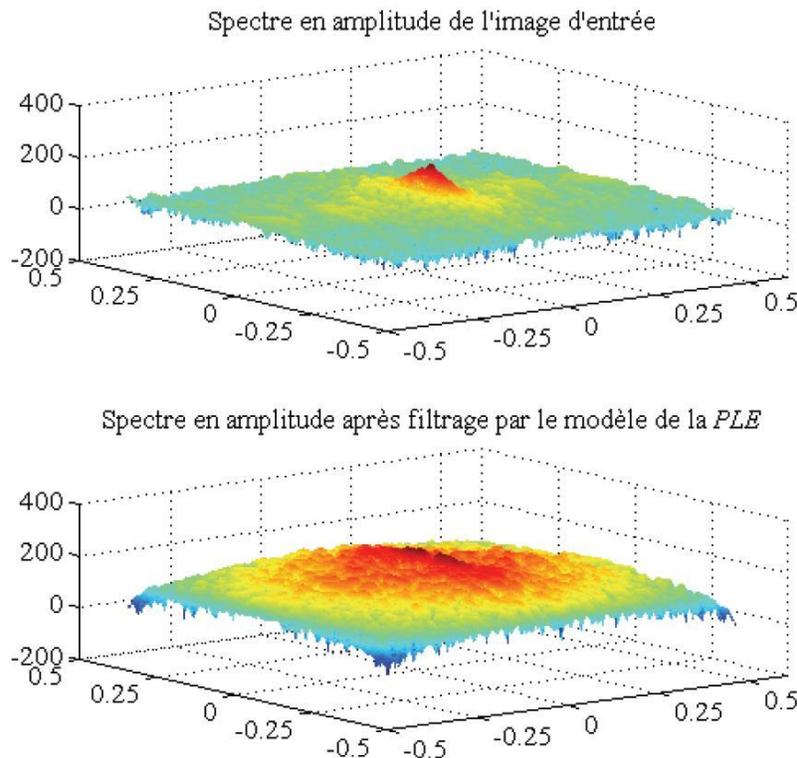
La **PLE** a aussi un effet de blanchiment spectral sur les images (*figure B.3*). Les hautes fréquences de l'image, qui sont associées aux détails, en particulier les contours sont rehaussés par le filtre modélisant la **PLE**. On distingue clairement sur le second spectre le rehaussement des hautes fréquences et une atténuation du mode principal de l'image filtrée par rapport au spectre de l'image principale où l'énergie est principalement concentrée dans les basses fréquences.

B.3 La Couche Plexiforme Interne

La couche plexiforme interne **PLI** est la dernière couche de la rétine avant la transmission de l'information visuelle vers le cerveau par le nerf optique. L'information provenant de la **PLE** est transmise par les cellules bipolaires aux cellules ganglionnaires et amacrines par l'intermédiaire de la **PLI**. La sortie de la **PLI** est constituée par les axones des cellules ganglionnaires qui forment le nerf optique. Le nerf optique véhicule l'information visuelle jusqu'aux aires visuelles du cerveau. Il existe une grande diversité de cellules ganglionnaires et amacrines dont les comportements ne sont pas encore modélisés avec précision.

On distingue 3 canaux résultant des interactions des différents types de cellules de la **PLI** : la voie Parvocellulaire qui est dédiée à la vision haute résolution et qui est sensible aux contrastes locaux dans l'image, la voie Magnocellulaire dédiée aux informations de mouvements et la voie Koniocellulaire dont le rôle n'est pas clairement défini à l'heure actuelle.

Du point de vue des images statiques, seul le canal Parvocellulaire, qui regroupe les traitements uniquement spatiaux de l'information visuelle, va nous intéresser. Le canal Parvocellulaire est constitué par les sorties des cellules ganglionnaires de type *P* qui agissent directement sur les

FIGURE B.3 – **Blanchiment Spectral** [Bouvier 10]

sorties des cellules bipolaires **ON** et **OFF**. Les cellules ganglionnaires adaptent leurs réponses aux sorties des cellules bipolaires d'une manière analogue aux photo-récepteurs, qui adaptent leur réponse à la luminance locale moyenne. Pour modéliser l'effet des cellules ganglionnaires de type *P* au niveau de la **PLI**, une compression adaptative est réalisée à la sortie des cellules bipolaires **ON** et **OFF** avec une loi de compression identique à celle utilisée pour les photo-récepteurs. Ensuite seulement, les sorties **ON** et **OFF** sont combinées pour obtenir la sortie **ON-OFF**.

B.4 Cortex *V1*

Une fois l'information passée à travers la *PLI*, les signaux des voies Parvocellulaire, Magnocellulaire et Koniocellulaire sont véhiculées par les nerfs optiques en direction du cerveau. L'information visuelle transite tout d'abord par les corps genouillés latéraux avant d'atteindre l'aire *V1* du cortex occipital. L'aire *V1* est la première étape du traitement de l'information visuelle par le cerveau. On dénombre environ une trentaine d'aires pouvant intervenir pendant le processus de vision mais seule l'aire *V1* est connu pour une modélisation précise. Le *SVH* procède à une analyse par bandes de fréquences et bandes d'orientations de l'information visuelle. C'est dans le cortex *V1* que ce phénomène a été mis en évidence.

Plusieurs approches ont été développées dont notamment les travaux de Guyader [Guyader 06] qui a introduit une modélisation du cortex *V1* par des filtres de Gabor Log Polaire **GLOP**. L'ajout d'une contrainte de symétrie sur l'axe des fréquences en échelle logarithmique permet de traiter les zooms et les changements d'échelles simplement et de modéliser des mesures spectrales proche de celle du système biologique.

Pour caractériser une image, la méthode consiste à échantillonner le spectre d'amplitude de l'image par une rosace de filtres **GLOP** dont l'expression est donnée par l'équation suivante :

$$G_{i,k}(f, \theta) = \frac{1}{2\pi\sigma^2} \frac{1}{f} \exp\left(-\frac{\ln(\frac{f}{f_k})^2}{2\sigma^2}\right) \exp\left(-\frac{(\theta - \theta_i)^2}{2\sigma^2}\right) \quad (\text{B.1})$$

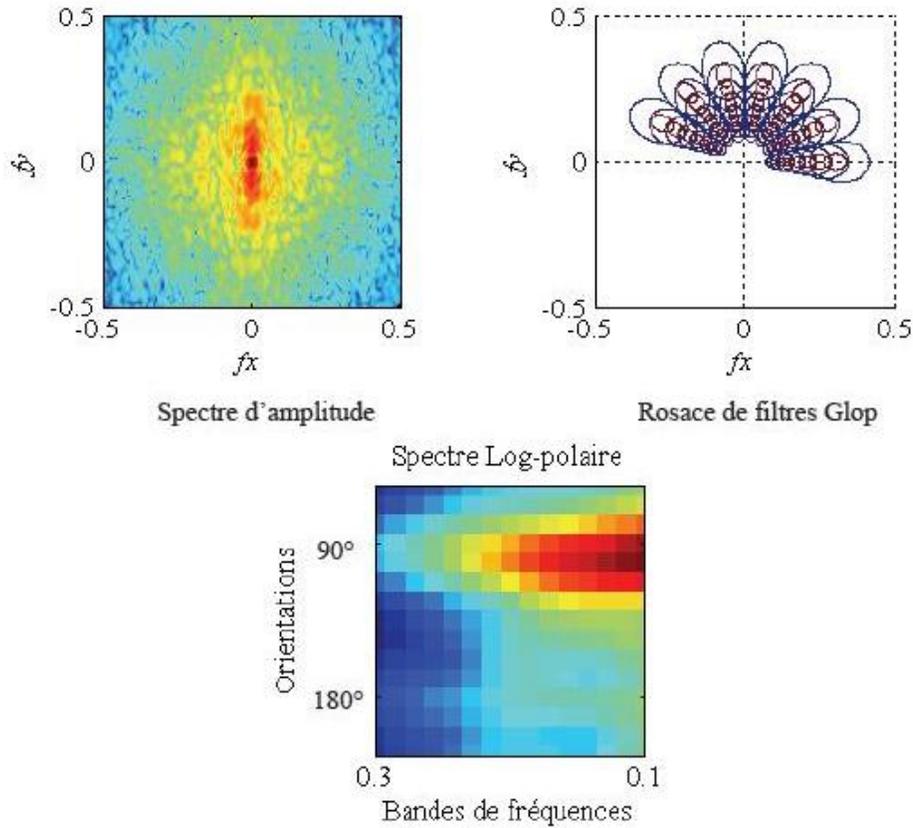


FIGURE B.4 – **Filtre Glop** [Bouvier 10]

C Les Espaces Couleurs



Sommaire

C.1	L'Espace RVB	203
C.2	Les Espaces de Télécommunication : $YCbCr$ - YUV - YIQ	204
C.3	Les Espaces Perceptuels HSV - HSL	205
C.4	Les Espaces Colorimétriques XYZ - La^*b^*	207

Nous présenterons ici une liste non exhaustive des espaces les plus couramment utilisés dans les applications de segmentations couleurs. Nous avons utilisé le logiciel ColorSpace¹ [Trémeau 03] qui effectue la conversion des principaux espaces couleurs et permet leurs visualisations selon 2 méthodes :

- une représentation *2D* : où chaque plan couleur est affiché avec des pseudo-teintes.
- une représentation *3D* : où les coordonnées de chaque couleur pour un espace couleur donné servent à déterminer leur position dans l'espace 3D.

Nous pouvons séparer les différents espaces couleurs en différentes catégories de par leur spécificité de représentation. Ainsi on peut identifier l'espace classique *RVB*, les espaces couleurs perceptuels calqués sur le système visuel humain, les espaces couleurs de télécommunications basés sur l'amélioration et le rendu numérique et enfin les espaces colorimétriques de référence.

C.1 L'Espace *RVB*

L'espace couleur *RVB* est l'espace de couleur le plus utilisé en colorimétrie numérique. En particulier, car il est à la base de nombreuses technologies numériques (écran, scanner, appareil photo...). Il est composé des trois couleurs primaires de la synthèse additive des couleurs : Rouge, Vert et Bleu. Elles correspondent aux trois longueurs d'ondes auxquels les cônes de l'oeil humain sont les plus sensibles. Il est souvent représenté comme un cube où chaque composante couleur représente une dimension.

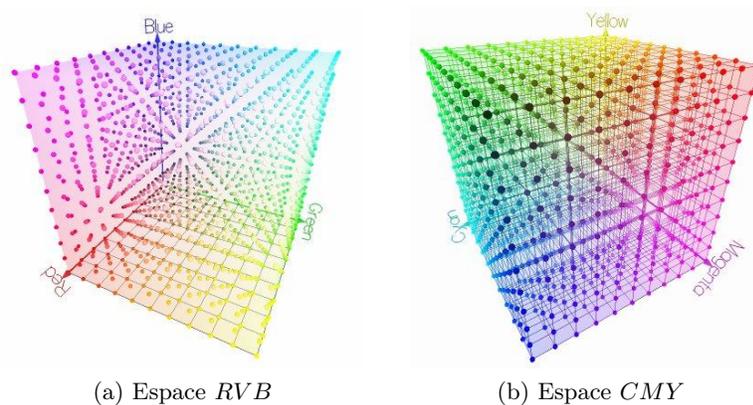


FIGURE C.1 – Les espaces *RVB* et *CMY*

On associe souvent à cet espace, l'espace *CMY* (Cian Magenta et Jaune (*Yellow*)) qui résulte de la synthèse soustractive des couleurs avec $C = 1 - R$, $M = 1 - V$ et $Y = 1 - B$.

L'espace *RVB* reste très utilisé de par sa grande simplicité malgré son inconvénient majeur qui est d'être un espace non uniforme entre ses différentes composantes qui le composent, ainsi les 'distances' couleurs ne sont pas linéaires avec le système visuel humain. Les composantes *RVB* sont fortement corrélées ce qui implique que les composantes de luminance et de chrominances ne soient pas séparées.

1. www.couleur.org

On trouve souvent aussi une représentation normalisée de l'espace RVB obtenue grâce aux équations : $r = \frac{R}{R+V+B}$ $v = \frac{V}{R+V+B}$ $b = \frac{B}{R+V+B}$

C.2 Les Espaces de Télécommunication : $YCbCr$ - YUV - YIQ

Les espaces $YCbCr$, YUV et YIQ sont issus des besoins technologiques pour les transmissions télévisuelles. Au début, la diffusion se faisait en noir et blanc et elle ne nécessitait que la composante de luminance. Avec l'évolution de la technologie, la télévision s'est vue dotée de la couleur et par conséquent de récepteurs couleurs. Néanmoins, il était indispensable de rester compatible avec l'ancien système et de respecter les deux points suivants : donner la possibilité de continuer à recevoir un signal en noir et blanc pour les postes de télévision en noir et blanc et de pouvoir visionner des images en noir et blanc sur un poste de télévision couleur.

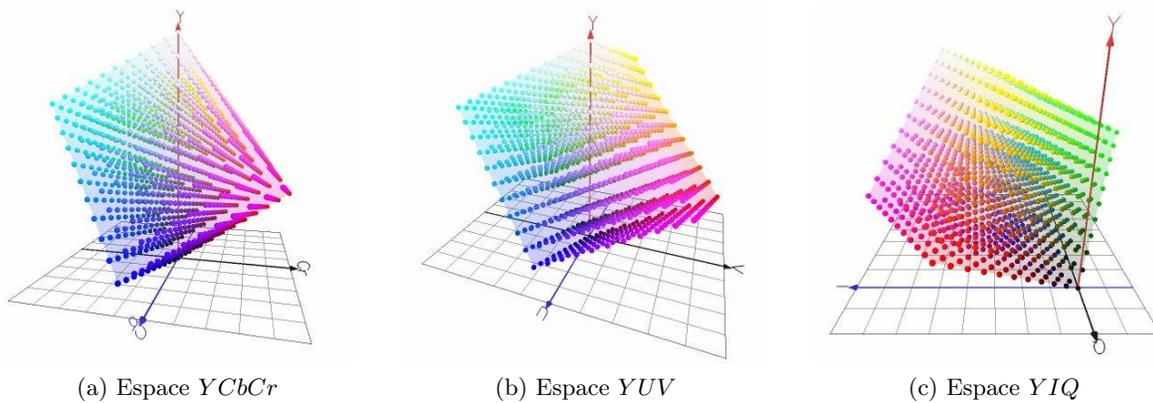


FIGURE C.2 – Les Espaces de Télécommunication

Ainsi, les signaux émis par les chaînes de télévision ont séparé l'information de luminance de celle de chrominance. L'espace YIQ a été introduit pour le standard NTSC pour la télévision américaine. L'espace YUV a été introduit pour le standard allemand PAL. Enfin l'espace $YCbCr$ est le standard international par excellence, dédié au codage digital des images de la télévision numérique et dans les formats de compression JPEG et MPEG.

Le principal intérêt de ces espaces est qu'il sépare la composante de Luminance (Y) des composantes de chrominances (Uv , IQ , $CbCr$). De plus l'intensité est encodée avec une précision supérieure que les composantes chromatiques afin de correspondre à l'acuité visuelle humaine et permet ainsi la compression de l'information couleur.

– L'espace YUV :

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{pmatrix} R \\ V \\ B \end{pmatrix} \quad (\text{C.1})$$

– L'espace YIQ :

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.212 & -0.523 & -0.311 \end{bmatrix} \cdot \begin{pmatrix} R \\ V \\ B \end{pmatrix} \quad (\text{C.2})$$

– L'espace YCbCr :

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{bmatrix} 0.2989 & 0.5866 & 0.1145 \\ -0.1688 & -0.3312 & 0.5 \\ 0.5 & -0.4184 & -0.0816 \end{bmatrix} \cdot \begin{pmatrix} R \\ V \\ B \end{pmatrix} \quad (\text{C.3})$$

C.3 Les Espaces Perceptuels HSV - HSL

Les espaces *HSV*, *HLS* sont des espaces créés à partir de l'analyse perceptuelle des couleurs basée sur le système visuel humain. Il représente ce que naturellement nous utilisons pour décrire une couleur : sa teinte (*H*), sa luminosité (*L* ou *V*) qui représente l'éclat d'une couleur ou la quantité de lumière réfléchiée par celle-ci et de son intensité (*S*) qui représente la saturation décrivant la pureté d'une teinte par rapport au blanc. Une couleur vive sera fortement saturée alors qu'une couleur pastel sera faiblement saturée.

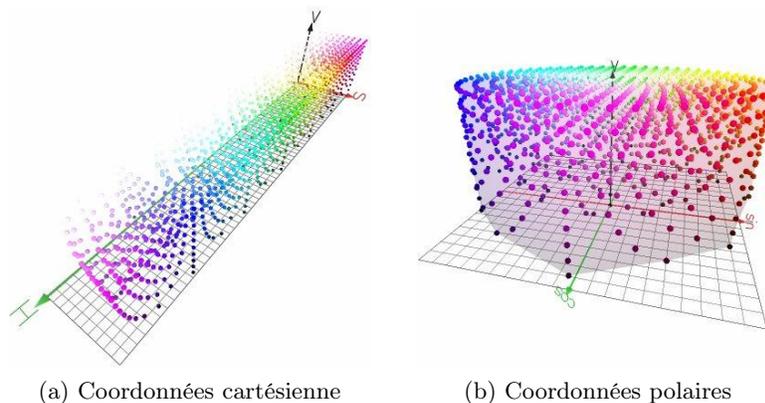
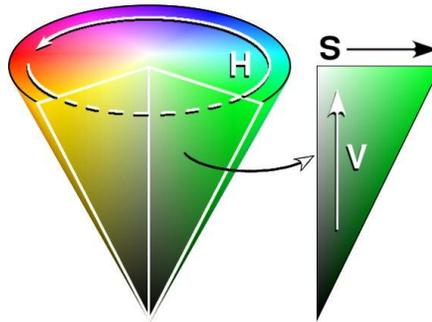


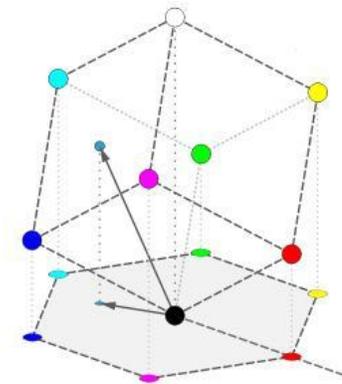
FIGURE C.3 – L'espace HSV

La représentation conique en coordonnées polaire est bien adaptée pour représenter chaque espace en un seul objet (*figure C.4a*). Ainsi, l'axe achromatique constitue l'axe des intensités (*L/V/I*) et la couleur sera définie par une position sur un palier circulaire où la saturation (*S*) représente le rayon et la teinte (*H*) représente l'angle.

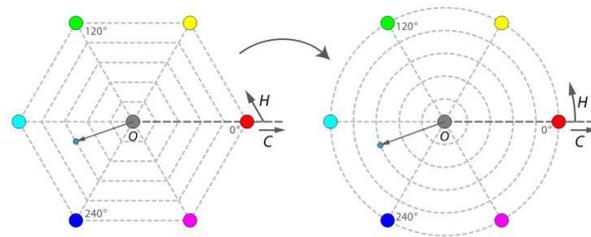
La valeur *H* de la teinte dérive d'une projection hexagonale du modèle *RVGB* (*figure C.4b*) que l'on adapte dans une représentation circulaire : chaque côté de l'hexagone est représenté par



(a) Représentation conique de L'espace HSV



(b) Projection du cube RGB en modèle Hexagonale HSV



(c) Passage de la représentation de la forme hexagonale à circulaire

FIGURE C.4 – L'espace HSV

un arc de cercle de 60° (figure C.4c).

$$H = 60^\circ \cdot \begin{cases} 0 & \text{si } M = m \\ \frac{V-B}{M-m} \text{ mod}(360^\circ) & \text{si } M = R \\ \frac{R-R}{M-m} + 120^\circ & \text{si } M = V \\ \frac{R-V}{M-m} + 140^\circ & \text{si } M = B \end{cases} \quad (\text{C.4})$$

avec $M = \max(R, V, B)$ et $m = \min(R, V, B)$.

La conversion en coordonnées polaires peut s'effectuer directement à partir des coordonnées cartésiennes, en dérivant la paire (α, β) des coordonnées cartésiennes de chromaticité :

$$\begin{aligned} H &= \arctan\left(\frac{\beta}{\alpha}\right) \\ \alpha &= R - \frac{1}{2}(G + B) \\ \beta &= \frac{\sqrt{3}}{2}(G - B) \end{aligned} \quad (\text{C.5})$$

À partir de la représentation de la composante H , on peut définir pour chaque espace couleur les équations des 2 autres composantes couleurs :

– L'espace *HSV* :

$$\begin{aligned} H &= H \\ S &= \begin{cases} 0 & \text{si } M = m \\ \frac{V-m}{V} & \text{sinon} \end{cases} \\ V &= M \end{aligned} \quad (\text{C.6})$$

– L'espace *HSL* :

$$\begin{aligned} H &= H \\ S &= \begin{cases} 0 & \text{si } M = m \\ \frac{M-m}{2L} & \text{si } L \leq \frac{1}{2} \\ \frac{M-m}{2-2L} & \text{si } L > \frac{1}{2} \end{cases} \\ L &= \frac{1}{2}(\max(R, V, B) + \min(R, V, B)) \end{aligned} \quad (\text{C.7})$$

– L'espace *HSI* (surtout utilisé pour des applications de vision par ordinateur) :

$$\begin{aligned} H &= H \\ S &= \begin{cases} 0 & \text{si } M = m \\ 1 - \frac{m}{M} & \text{sinon} \end{cases} \\ I &= (R + V + B)/3 \end{aligned} \quad (\text{C.8})$$

C.4 Les Espaces Colorimétriques XYZ - La*b*

Les espaces colorimétriques sont définis par la *Commission internationale de l'éclairage* (CIE).

L'espace CIE XYZ est le premier espace créé en 1931. En considérant une lumière quelconque décomposée en une somme de lumières de couleurs pures dont les intensités peuvent être multipliées par leurs fonctions d'efficacité. On obtient un résultat *Y* appelé luminance. Celle-ci s'exprime en fonction des composantes Rouge, Verte et Bleue par l'équation :

$$Y = 0,299R + 0,587V + 0,114B \quad (\text{C.9})$$

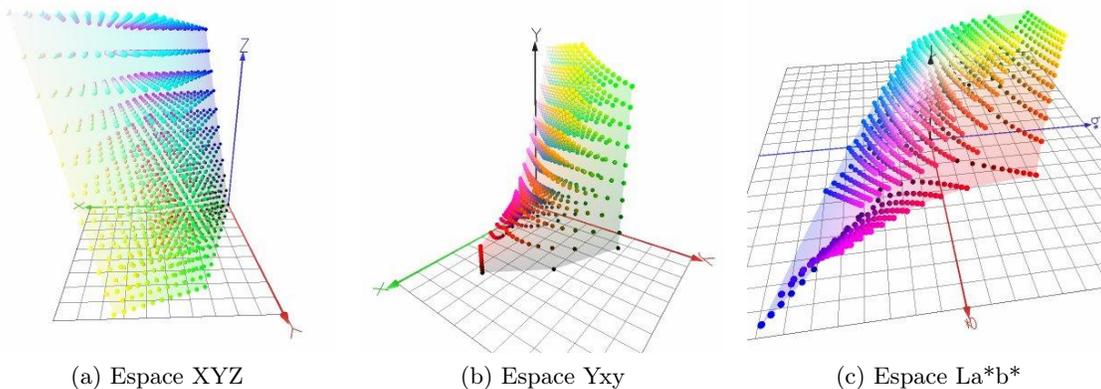


FIGURE C.5 – Les Espaces Colorimétriques

À la grandeur Y , la CIE a ajouté deux autres grandeurs indépendantes notées X et Z pour définir un système équivalent au RVB.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix} \cdot \begin{pmatrix} R \\ V \\ B \end{pmatrix} \quad (\text{C.10})$$

avec les coefficients de la matrice de passage conditionnés par le blanc de référence.

À partir de l'espace XYZ , on peut définir l'espace CIE Yxy . L'idée de base de ce système consiste à associer la luminance Y à deux valeurs (x, y) qui définissent la chrominance indépendamment de l'intensité dans le plan d'équation $X + Y + Z = 1$ projeté dans le plan (XY) . On définit alors le diagramme de chromaticité en 'fer à Cheval' représentant la teinte et la saturation de l'ensemble des couleurs (figure C.6).

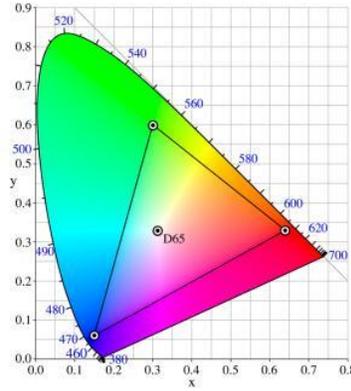


FIGURE C.6 – **L'espace CIE xy** Représentation du diagramme de chromaticité auquel on a inclus la dynamique de l'espace RGB (triangle noir) sous illumination naturelle D65

À partir de l'espace XYZ , on peut aussi définir l'espace CIE La^*b^* qui essaye de modéliser la réponse logarithmique de l'oeil. En outre, il possède le grand avantage d'être uniforme. Le modèle de couleur La^*b^* a été créé comme un modèle absolu, indépendant du matériel, utilisable comme référence théorique.

$$L = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \quad \text{si} \quad \frac{Y}{Y_n} > 0.008856 \quad (\text{C.11})$$

$$L = 903.3 \quad \text{sinon} \quad (\text{C.12})$$

$$a = 500\left(f\left(\frac{X}{X_n}\right) - \left(\frac{Y}{Y_n}\right)\right) \quad (\text{C.13})$$

$$b = 200\left(f\left(\frac{Y}{Y_n}\right) - \left(\frac{Z}{Z_n}\right)\right) \quad (\text{C.14})$$

où

$$f(t) = (t)^{\frac{1}{3}} \quad \text{si} \quad t > 0.008856 \quad (\text{C.15})$$

$$f(t) = 7.7787t + \frac{16}{116} \quad \text{sinon} \quad (\text{C.16})$$

avec X_n, Y_n, Z_n les coordonnées du blanc de référence qui est choisi selon la nature de l'illuminant.



(a) Image Originale



(b) R



(c) V



(d) B



(e) Hue



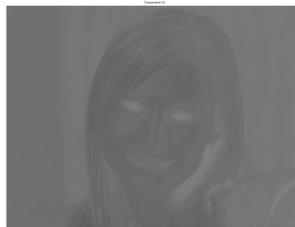
(f) Saturation



(g) Value



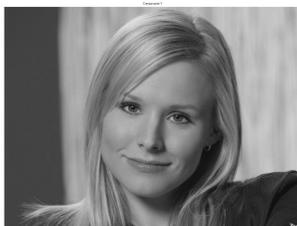
(h) Y



(i) Cb



(j) Cr



(k) L



(l) a*



(m) b*

FIGURE C.7 – Décomposition des Espaces Couleur : Représentation de chaque composante des principaux espaces couleurs

Bibliographie

- [ARb] *The AR face database.* <http://www2.ece.ohio-state.edu/aleix/ARdatabase.html>.
- [Auzou 01] M.C. Auzou & S.Melchior-Bonnet. *Les vies du cheveu.* Gallimard, 2001.
- [Beaudot 94] W. Beaudot. *Le traitement neuronal de l'information dans la rétine des vertébrés : Un creuset d'idées pour la vision artificielle.* Master's thesis, Institut National Polytechnique de Grenoble - INPG, 1994.
- [BeauteFamille] BeauteFamille. <http://beaute.notrefamille.com/>.
- [Bellot 02] D. Bellot. *Fusion de données avec des réseaux bayésiens pour la modélisation des systèmes dynamiques et son application en télémédecine.* Master's thesis, Université Henri Poincaré - Nancy 1, 2002.
- [Benoit 07] A. Benoit. *Le système visuel humain au secours de la vision par ordinateur.* PhD thesis, Institut National Polytechnique de Grenoble - INPG, 2007.
- [Berman 00] A. Berman, P. Vlahos & A. Dadourian. *Method for removing from an image the background surrounding a selected object.* US patent 6,135,345, 2000.
- [Bertails 06] F. Bertails. *Simulation de chevelures virtuelles.* Master's thesis, Institut National Polytechnique de Grenoble - INPG, 2006.
- [Beucher 79] S. Beucher & C. Lantuejoul. *Use of watersheds in contour detection.* International Workshop on Image Processing : Real-time Edge and Motion Detection/Estimation, 1979.
- [Beucher 92] S. Beucher. *The watershed transformation applied to image segmentation.* Conf. on Signal and Image Processing in Microscopy and Microanalysis, vol. 6, pages 299–314, 1992.
- [Blakemore 69] C. Blakemore & F.W. Campbell. *On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images.* Journal of Physiology, vol. 203, pages 237–260, 1969.
- [Bouvier 10] C. Bouvier. *Segmentation region-contour des contours des levres.* Master's thesis, Institut National Polytechnique de Grenoble - INPG, Février 2010.
- [Boykov 01] Y.Y. Boykov & M.P. Jolly. *Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images,* 2001.
- [Canny 86] J. Canny. *A computational approach to edge detection.* Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pages 679–698, 1986.
- [Chai 98] D. Chai & K.N. Ngan. *Locating Facial Region of a Head-and-Shoulders Color Image.* International Conference on Faceand Gesture Recognition, pages 124–129, 1998.

- [Chai 99] D. Chai & K.N. Ngan. *Face segmentation using Skin color map in Video-phone Applications*. IEEE transactions on Circuits and Systems for Video Technology, vol. 9, no. 4, pages 551–564, 1999.
- [Chen 06] H. Chen & S.C. Zhu. *A Generative Sketch Model for Human Hair Analysis and Synthesis*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, pages 1025–1040, 2006.
- [Cho 07] T.S. Cho, W.T. Freeman & H. Tsao. *A reliable skin mole localization scheme*. ICCV, pages 1–8, 2007.
- [Chuang 01] Y.Y. Chuang, B. Curless, D.H. Salesin & R. Szeliski. *A Bayesian Approach to Digital Matting*. Computer Vision and Pattern Recognition, vol. 2, pages 264–271, 2001.
- [Coggins 83] J.M. Coggins. *A framework for texture analysis based on spatial filtering*. PhD thesis, Michigan State University, 1983.
- [Coiffeursdefrance] Coiffeursdefrance. <http://www.coiffeursdefrance.com/>.
- [Comaniciu 02] D. Comaniciu & P. Meer. *Mean Shift : A robust approach toward feature space analysis*. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pages 603–619, 2002.
- [Daubechies 92] I. Daubechies. *Ten lectures on wavelets*. SIAM, 1992.
- [Dempster 77] A. P. Dempster, N. M. Laird & D. B. Rubin. *Maximum likelihood from incomplete data via the em algorithm*. Journal of the Royal Statistical Society, vol. B, no. 39, pages 1–38, 1977.
- [Deriche 87] R. Deriche. *Using canny’s criteria to derive an optimal edge detector recursively implemented*. International Journal on Computer Vision. International Journal on Computer Vision, vol. 2, pages 15–20, 1987.
- [Dubois 88] D. Dubois & H. Prade. *Possibility theory*. Plenum Press, New-York, 1988.
- [Dubois 92] D. Dubois & H. Prade. *Putting rough sets and fuzzy sets together*. In R Slowinski, editeur, *Intelligent Decision Support- Handbook of Applications and Advances of the Rough Set Theory*, pages 203–232. Kluwer Academic, Dordrecht, 1992.
- [Durette 05] B. Durette & J. Hérault. *Traitement visuels biomimétiques pour la suppléance perceptive*. Rapport technique, Gipsa-lab, 2005.
- [Elgammal 09] A. Elgammal, C. Muang & D. Hu. *Skin Detection, a Short Tutorial*. Encyclopedia of Biometrics, 2009.
- [Gemey] Gemey. <http://www.gemey-maybelline.com/>.
- [Germain 03] C. Germain, J.P. Da Costa, O. Laviolle & P. Baylou. *Multiscale estimation of vector field anisotropy application to texture characterization*. Signal Processing, vol. 83, no. 7, pages 1487–1503, 2003.
- [Guyader 06] N. Guyader, A. Chauvin, C. Massot, J. Héraul & C. Marendaz. *A biological model of low-level vision suitable for image analysis and cognitive visual perception*. European Conference on Visual Perception, vol. 35, 2006.
- [Hall 01] D.L. Hall & J.Llinas. *Handbook of multisensor data fusion*. Numéro 1. CRC Press, 2001.
- [Haralick 79] R.M. Haralick. *Statistical and Structural Approaches to Texture*. Proceedings of the IEEE, vol. 67, no. 786-804, 1979.

- [Harvey 90] L.O. Harvey & V.V. Doan. *Visual masking at different polar angles in the two dimensional Fourier plane*. Journal of Optical Society of America, vol. 7, no. 116-127, 1990.
- [Harville 05] M. Harville, H. Baker, N. Bhatti & S. Susstrunk. *Consistent Image-Based Measurement and Classification of Skin Color*. International Conference on Image Processing, pages 11–14, 2005.
- [Horprasert 99] T. Horprasert, D. Harwood & L.S. Davis. *A statistical approach for real-time robust background subtraction and shadow detection*. pages 1–19, 1999.
- [Hérault 01] J. Hérault. *De la rétine biologique aux circuits neuromorphiques*. In Hermès, editeur, Les Systèmes de Vision, 2001.
- [Hsu 02] R. L. Hsu, M. Abdel-Mottaleb & A.K. Jain. *Face detection in color images*. IEEE Trans. on PAMI, vol. 24, no. 5, pages 696–706, 2002.
- [Hubel 62] D.H. Hubel & T.N. Wiesel. *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex*. Journal of Physiology, vol. 160, pages 106–154, 1962.
- [Jarudi 07] I.N. Jarudi & P. Sinha. *Relative Contributions of Internal and External Features to Face Recognition*. Rapport technique, Defense Technical Information Center OAI-PMH Repository, 2007.
- [Jedynak 03] B. Jedynak, Huicheng Zheng & Mohamed Daoudi. *Statistical Models for Skin Detection*. In Statistical Analysis in Computer Vision, 2003.
- [Jones 99] Michael J. Jones & J. M. Rehg. *Statistical Color Models with Application to Skin Detection*. International Journal of Computer Vision, pages 274–280, 1999.
- [Jones 04] M.J. Jones & J.M. Rehg. *Statistical Color Models with Application to Skin Detection*. International Journal of Computer Vision, vol. 46, no. 1, pages 81–96, 2004.
- [Julesz 62] B. Julesz. *Visual Pattern Discrimination*. Transactions on Information Theory, vol. 8, pages 84–92, 1962.
- [Julesz 87] B. Julesz, & J. R. Bergen. *Textons, the fundamental elements in preattentive vision and perception of textures*. In Readings in computer vision : issues, problems, principles, and paradigms, pages 243–256, 1987.
- [Julesz 95] B. Julesz. *Dialogues on perception*. Cambridge : Bradford/MIT Press, 1995.
- [Kajiya 89] J.T. Kajiya & T. L. Kay. *Rendering fur with three dimensional textures*. SIGGRAPH, vol. 23, no. 3, pages 271–280, 1989.
- [Kakumanu 06] P. Kakumanu, S. Makrogiannis & N. Bourbakis. *A survey of skin-color modeling and detection methods*. Pattern Recognition, vol. 40, no. 3, pages 1106–1122, 2006.
- [Kampmann 97a] M. Kampmann. *Estimation of the chin and cheek contours for precise face model adaptation*. Image Processing, 1997. Proceedings., International Conference on, vol. 3, pages 300–303, 1997.
- [Kampmann 97b] M. Kampmann & J. Ostermann. *Automatic adaptation of a face model in a layered coder with an object-based analysis-synthesis layer and a knowledge-based layer*. Signal Processing : Image Communication, vol. 9, no. 3, pages 201–220, 1997.

- [Kampmann 98] M. Kampmann. *Segmentation of a head into face, ears, neck and hair for knowledge-based analysis-synthesis coding of videophone sequences*. Image Processing, International Conference on, vol. 2, page 876, 1998.
- [Karlsson 04] F. Karlsson. *Matting of natural image sequences using bayesian statistics*. Master's thesis, Linköpings Universitet (Suède), 2004.
- [Kovac 03] J. Kovac, P. Peer & F. Solina. *Human Skin Colour Clustering for Face Detection*. International Conference on Computer, 2003.
- [Lee 08] K.C. Lee, D. Anguelov, B. Sumengen & S.B. Gokturk. *Markov random field models for hair and face segmentation*. FG08, pages 1–8, 2008.
- [Levin 04] A. Levin, D. Lischinski & Y. Weiss. *Colorization using optimization*. ACM Trans. Graph., vol. 23, no. 3, pages 689–694, 2004.
- [Levin 06] A. Levin, D. Lischinski & Y. Weiss. *A Closed Form Solution to Natural Image Matting*. CVPR, vol. 1, pages 61–68, 2006.
- [Levin 08] A. Levin, D. Lischinski & Y. Weiss. *A Closed-Form Solution to Natural Image Matting*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 2, pages 228–242, 2008.
- [Lienhart 03] R. Lienhart, L. Liang & E. Kuranov. *A detector tree of boosted classifiers for real-time object detection and tracking*. In IEEE ICME2003, pages 277–280, 2003.
- [Longère 98] P. Longère & A. Tremeau. *Color Appearance : Effects of Texture and Relief*. In Electronic Imaging, pages 89–97, 1998.
- [LOreal] L'Oréal. <http://www.loreal-paris.fr/mon-maquillage-virtuel.aspx>.
- [Makeover] Makeover. <http://makeover.dailymakeover.com/virtual-makeover/>.
- [Mallat 01] S. Mallat. *Une exploration des signaux en ondelettes*. Ecole Polytechnique, 2001.
- [Mandelbrot 75] B. Mandelbrot. *Les objets fractals : forme, hasard et dimension*. Flammarion, 1975.
- [Marschner 03] S.R. Marschner, H.W. Jensen, M. Cammarano, S. Worley & P. Hanrahan. *Light scattering from human hair fibers*. ACM Transactions on Graphics, vol. 22, pages 780–791, 2003.
- [McNeil 94] J.E. McNeil, L. Cipolotti & E.K. Warrington. *The accessibility of proper names*. Neuropsychologia, vol. 32, no. 2, pages 193–208, 1994.
- [Mercier 08] D. Mercier, B. Quost & T. Denoex. *Refined modeling of sensor reliability in the belief function framework using contextual discounting*. Information Fusion, vol. 9, no. 2, pages 246–258, 2008.
- [Michelet 06] F. Michelet, J.P. Da Costa, P. Baylou & C. Germain. *Local Orientation Estimation in Corrupted Images*. IWICPAS, pages 349–358, 2006.
- [Michelet 07] F. Michelet, J.P. Da Costa, O. Lavialle, Y. Berthoumieu, P. Baylou & C. Germain. *Estimating local multiple orientations*. Signal Processing, vol. 87, no. 7, pages 1655–1669, 2007.
- [Porter 84] T. Porter & T. Duff. *Compositing Digital Images*. SIGGRAPH, vol. 18, no. 3, pages 253–259, 1984.

- [Rammasso 07] E. Rammasso. Reconnaissance de séquences d'états par le modèle des croyances transférables : Application à l'analyse de vidéos d'athlétisme. Master's thesis, Institut National Polytechnique de Grenoble - INPG, Décembre 2007.
- [Robbins 02] C. Robbins. Chemical and physical behavior of human hair. Springer, 2002.
- [Rousset 08] C. Rousset & P.Y. Coulon. *Frequential and Color Analysis for Hair Mask Segmentation*. ICIP, pages 2276–2279, 2008.
- [Russ 94] J.C. Russ. Fractal surface. Plenum, 1994.
- [Ruzon 00] M.A. Ruzon & C. Tomasi. *Alpha Estimation in Natural Images*. Computer Vision and Pattern Recognition, vol. 1, pages 1018–1025, 2000.
- [Science] L'Oréal Hair Science. <http://www.hair-science.fr>.
- [Shafer 76] G. Shafer. A mathematical theory of evidence. Princeton University Press, 1976.
- [Shin 02] M.C. Shin, K.I. Chang & L.V. Tsap. *Does Colorspace Transformation Make Any Difference on Skin Detection ?* Applications of Computer Vision, IEEE Workshop on, page 275, 2002.
- [Sinha 96] P. Sinha & T. Poggio. *I think I know that face...* Nature, vol. 384, page 404, Dec 1996. Provided by the SAO/NASA Astrophysics Data System.
- [Smets 93] P. Smets. *Beliefs functions : The Disjunctive Rule of Combination and the Generalized Bayesian Theorem*. Approximate Reasoning, vol. 9, pages 1–35, 1993.
- [Smets 94] P. Smets & R. Kennes. *The transferable belief model*. vol. 66, no. 2, pages 191–234, 1994.
- [Smets 05] P. Smets. *Decision making in the TBM : the necessity of the pignistic transformation*. Int. J. Approx. Reasoning, vol. 38, no. 2, pages 133–147, 2005.
- [Smith 96] A.R. Smith & J.F. Blinn. *Blue Screen Matting*. Microsoft Corporation, 1996.
- [Sobottka 96] K. Sobottka & I. Pitas. *Face localization and feature extraction based on shape and color information*. International conference on Image Processing, pages 483–486, 1996.
- [Stillitano 09] S. Stillitano. Segmentation et suivi des contours externe et interne des lèvres pour des applications de maquillage virtuel et de labiophonie. Master's thesis, Institut National Polytechnique de Grenoble - INPG, Mai 2009.
- [Sun 06] J. Sun, Y. Li, S.B. Kang & H.Y. Shum. *Flash matting*. ACM Trans. Graph., vol. 25, no. 3, pages 772–778, 2006.
- [Sun 07] J. Sun, Y. Li, S.B. Kang & H.Y. Shum. *Flash Cut : Foreground Extraction with Flash and No-flash Image Pairs*. CVPR, 2007.
- [Taa] <http://www.taaz.com/>.
- [Tang 06] C.K. Tang, J. Sun, J. Jia & H.Y. Shum. *Poisson Matting*. Technologies of E-learning and Digital Entertainment, 2006.
- [Torres 01] L. Torres, A. Albiol & E.J. Delp. *An unsupervised color image segmentation algorithm for face detection applications*. International Conference on Image Processing, vol. 2, pages 681–684, 2001.

- [Trémeau 97] A. Trémeau & N. Borel. *A region growing and merging algorithm to color segmentation*. Pattern Recognition, vol. 30, pages 1191–1203, 1997.
- [Trémeau 03] A. Trémeau & P. Colantoni. *Color data visualization for color imaging*. VCIP, pages 1068–1077, 2003.
- [Tuceryan 98] M. Tuceryan & A.K. Jain. The handbook of pattern recognition and computer vision, pages 207–248. World Scientific Publishing Co, 1998.
- [Valois 82] R. De Valois. *The orientation and direction selectivity of cells in macaque visual cortex*. Vision Research, vol. 22, pages 531–544, 1982.
- [Vesalis] Vesalis. <http://www.vesalis.com/>.
- [Vincent 91] L. Vincent & P. Soille. *Watersheds in Digital Spaces : An Efficient Algorithm Based on Immersion Simulations*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, pages 583–598, 1991.
- [Viola 01] P. Viola & M.J. Jones. *Rapid object detection using a boosted cascade of simple features*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
- [Viola 04] P. Viola & M.J. Jones. *Robust Real-Time Face Detection*. International Journal of Computer Vision, vol. 57, no. 2, 137-154 2004.
- [Wang 09] D. Wang, S.G. Shan, W. Zeng, H.M. Zhang & X.L. Chen. *A novel two-tier Bayesian based method for hair segmentation*. ICIP, pages 2401–2404, 2009.
- [Yacoob 06] Y. Yacoob & L.S. Davis. *Detection and Analysis of Hair*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 7, pages 1164–1169, 2006.
- [Yager 87] R.R. Yager. *On the dempster-shafer framework and new combination rules*. Inf. Sci., vol. 41, no. 2, pages 93–137, 1987.
- [Yang 99] M.H. Yang & N. Ahuja. *Gaussian Mixture Model for Human Skin Color and Its Applications in Image and Video Databases*. Proceedings of SPIE, pages 458–466, 1999.
- [Yedidi 03] J.S. Yedidi, W.T. Freeman & Y. Weiss. *Understanding belief propagation and its generalizations*. In Exploring artificial intelligence in the new millennium, pages 239–269, 2003.
- [Zadeh 65] L.A. Zadeh. *Fuzzy Sets*. Information Control, vol. 8, pages 338–353, 1965.
- [Zhang 05] J. Zhang, P. Regtien & M. Korsten. *Monitoring of a dry sliding wear using fractal analysis*. Metrology and Measurement Systems, vol. 12, no. 2, pages 111–118, 2005.
- [Zhang 08] Z. Zhang, H. Gunes & M. Piccardi. *An accurate algorithm for head detection based on XYZ and HSV hair and skin color models*. ICIP08, pages 1644–1647, 2008.
- [Zhang 09] Z. Zhang, H. Gunes & M. Piccardi. *Head Detection For Video Surveillance Based on Categorical Hair and Skin Colour Model*. ICIP09, pages 1137–1140, 2009.