

THÈSE

pour l'obtention du Grade de
DOCTEUR DE L'UNIVERSITÉ DE POITIERS
(Université de Poitiers)
(Diplôme National - Arrêté du 25 Avril 2002)

École Doctorale : **Sciences Pour l'Ingénieur et Aéronautique (SPI & A)**
Secteur de Recherche : **Traitement du Signal et des Images**

Présentée par :
Michèle GOUIFFÈS

Apports de la Couleur et des Modèles de Réflexion pour l'Extraction et le Suivi de Primitives

Directeur de Thèse :
Christine FERNANDEZ

soutenue le 1 Décembre 2005 devant la Commission d'Examen composée de :

Michel Dhome , Rapporteur
Bernard Péroche , Rapporteur
Christophe Collewet , Examineur
Sylvie Philipp-Foliguet , Examineur
Alain Trémeau , Co-directeur de Thèse
Christine Fernandez , Directeur de Thèse
Theo Gevers , Invité

REMERCIEMENTS

Les travaux présentés dans ce manuscrit ont été réalisés au Cemagref de Rennes, en collaboration avec le laboratoire SIC de l'Université de Poitiers et le laboratoire LIGIV de St-Etienne. Ils ont été financés par le Cemagref et la région Poitou-Charentes. Sans le soutien et la présence de nombreuses personnes, ce projet n'aurait pu aboutir.

Tout d'abord, j'adresse toute ma reconnaissance à Monsieur Petit, directeur du Cemagref de Rennes, pour m'avoir accueillie au sein de ce laboratoire.

Je tiens également à adresser mes plus sincères remerciements à Pascale Marty-Mahé, qui est à l'origine de cette thèse et de son financement. Je la remercie de m'avoir encadrée pendant la première année. J'ai apprécié sa sympathie et regretté son départ et je lui souhaite à présent d'être heureuse dans sa nouvelle carrière.

Merci énormément à Christophe Collewet qui a accepté de poursuivre mon encadrement de thèse. Il m'a proposé de nouvelles orientations et m'a permis d'élargir mes connaissances. Son perfectionnisme, ses conseils éclairés, sa grande disponibilité et sa grande sympathie ont été pour moi un soutien immense et un exemple à suivre.

Je remercie Christine Fernandez-Maloigne, ma directrice de thèse, pour sa disponibilité malgré l'éloignement, pour sa sympathie, son enthousiasme, son implication dans la réussite de ces travaux, ses conseils administratifs et ses encouragements tout au long de cette thèse.

Merci également à Alain Trémeau, mon co-directeur de thèse, pour ses conseils, son soutien, son implication au cours de cette thèse, mais également pour sa bonne humeur. Je tiens à le remercier aussi pour m'avoir accueillie et aidée quelques jours au LIGIV pour effectuer mes expérimentations, malgré leur aspect « odorant » !

J'exprime toute ma gratitude et mon respect aux membres du jury qui malgré un emploi du temps serré, m'ont tous fait l'honneur d'accepter de prendre part à ce jury et surtout de lire et d'expertiser mon travail. Je remercie Madame Sylvie Philipp-Foliguet, Professeur à l'ENSEA de Cergy Pontoise, qui a accepté la présidence du jury. Merci à Monsieur Bernard Péroche, Professeur à l'université de Lyon I, et à Monsieur Michel Dhome, directeur de recherche au LASMEA de Clermont Ferrand, qui ont accepté d'être rapporteurs de cette thèse, et qui l'ont fait

avec une ardeur et une minutie qui sont pour moi une récompense. Merci à Monsieur Theo Gevers, Professeur à l'Université d'Amsterdam, d'avoir effectué ce long déplacement pour assister à ma soutenance, et de s'être intéressé au manuscrit, pourtant écrit en français. Enfin, merci à tous pour toutes ces remarques pertinentes et ces questions intéressantes. Ce fut pour moi une journée très enrichissante.

Un grand merci également à toutes les personnes avec qui j'ai partagé le bureau, stagiaires, doctorants, post-doctorants : Ali, Dimitry, Anne-Claire, Jordi, Marc, Véronique, Jean-Marie. Ca a été pour moi de réels moments de plaisir. Merci à toutes les personnes du Cemagref, que j'ai cotoyées de près ou de loin, que ce soit à la pause café, à la pause déjeuner, pendant des footings, pendant des soirées. Merci pour cette très bonne ambiance qui régnait dans ces locaux. Un grand merci également à toutes les personnes du laboratoire SIC, pour leur accueil et leur bonne humeur, et pour tous les bons moments que je passe avec eux depuis deux mois.

Je remercie vivement ma famille, ma « belle-famille » et mes amis pour leur soutien et leurs encouragements, mais aussi tout simplement pour tous les bons moments que j'ai passés avec eux, pour tous ces éclats de rire, pour toutes ces soirées, autour d'une table, d'un mah jong, à l'Appatame, au théâtre, en concert, au cinéma ou bien encore dans un atelier de dessin, etc. En bref, merci de m'avoir permis de voyager de temps en temps à dix mille lieues du boulot.

Enfin, merci à Anthony, mon plus grand soutien au cours de ces trois années de thèse.

Table des matières

Introduction générale	1
I L'interaction lumière-matière-observateur	7
Introduction	9
1 L'interaction Lumière-Matière	11
1.1 La physique de l'interaction lumière-matière	12
1.1.1 La lumière	12
1.1.2 La matière	14
1.1.3 La reflectance : une mesure de l'interaction lumière-matière	16
1.2 Modélisation des phénomènes d'interaction lumière-matière	18
1.2.1 Les modèles de rugosité de surface	19
1.2.2 Les modèles de réflexion	21
1.3 Conclusion	24
2 La luminance et la couleur	27
2.1 La luminance	28
2.1.1 Luminance mesurée par un capteur CCD	28
2.1.2 Luminance perçue par l'œil humain	28
2.2 La couleur	29
2.3 La représentation de la couleur	32
2.3.1 Les systèmes de primaires	33
2.3.2 Les systèmes perceptuellement uniformes	36
2.3.3 Systèmes perceptuels	38
2.4 Systèmes d'axes indépendants : système de Ohta	39
2.4.1 Choix du système couleur	40
2.5 Conclusion	41
Conclusion	43

II	Apport de la couleur pour l'extraction de primitives	45
	Introduction	47
3	Extraction de primitives couleur : état de l'art en segmentation d'images couleur	49
3.1	L'approche contour	50
3.1.1	Approches marginales	51
3.1.2	Approche multicomposante ou vectorielle	52
3.1.3	Discussion	54
3.2	La classification pixellaire	55
3.2.1	Le <i>clustering</i>	55
3.2.2	La classification par analyse d'histogrammes	58
3.2.3	Classification spatio-colorimétrique	60
3.2.4	Discussion	62
3.3	Segmentation dans l'espace géométrique de l'image	63
3.3.1	Approche ascendante ou croissance de région	63
3.3.2	Approche structurelle : division-fusion de régions	64
3.3.3	Discussion	65
3.4	Méthodes mixtes	66
3.4.1	L'association d'approches de type région et de type contour	66
3.4.2	Approches classification-région	67
3.5	Discussions et conclusions	68
4	Détection d'un marquage à l'encre	71
4.1	La transparence	72
4.2	Étude de la couleur de l'encre imprimée sur un support de couleur uniforme	73
4.2.1	<i>Locus</i> des couleurs associées au marquage à l'encre	74
4.2.2	Les coefficients de concentration	76
4.2.3	Les rapports des coefficients de concentration marginaux	80
4.3	Algorithme de segmentation d'un marquage à l'encre	81
4.3.1	Remarques préliminaires	82
4.3.2	Sélection des pixels candidats	83
4.3.3	Élimination des faux candidats par calcul d'un score d'appartenance au marquage	85

4.3.4	Sélection d'une composante connexe associée au marquage	87
4.4	Exemples de segmentation	90
4.5	Conclusion	93
5	Application au contrôle de la traçabilité porcine	99
5.1	Système de vision et base d'images	101
5.2	Approches de segmentation par utilisation des invariants dédiés à l'encre	102
5.2.1	Extraction des ombres à l'aide des rapports de coefficients de concentration	103
5.2.2	Segmentation avec analyse de voisinage	104
5.2.3	Segmentation sans analyse de voisinage	105
5.3	Segmentation dans un espace couleur classique	105
5.3.1	Classification spatio-colorimétrique	107
5.3.2	Reconnaissance du marquage par analyse de texture	109
5.4	Résultats expérimentaux	110
5.5	Conclusion	114
	Conclusion	119
III	Apport de la couleur et des modèles de réflexion pour le suivi de primitives	121
	Introduction	123
6	Modélisation des variations photométriques dans des séquences d'images	125
6.1	Variations de luminance et de couleur	126
6.1.1	Variations de spécularité	128
6.1.2	Variations d'éclairage et de spécularité	129
6.2	Modèles locaux de variation d'illumination	131
6.2.1	Modèle de conservation de la luminance	131
6.2.2	Le modèle additif	132
6.2.3	Le modèle scalaire	132
6.2.4	Le modèle affine	132

6.3	Modèles de réflexion locaux adaptés à l'occurrence de spéularités et aux variations d'éclairage	133
6.3.1	Modèle d'illumination local	133
6.3.2	Modèle d'illumination complet	135
6.4	Validité des modèles locaux	136
6.4.1	Modélisation de la géométrie de la scène	136
6.4.2	Étude de la validité de la fonction λ	137
6.4.3	Étude de la validité de la fonction η	140
6.5	Conclusion	145
7	Suivi de primitives robuste aux variations d'illumination dans les images de lumiance	147
7.1	Détection de points d'intérêt	148
7.1.1	Méthodes de détection applicables aux images structurées	148
7.1.2	Méthodes de détection applicables aux images texturées	149
7.2	Suivi de points d'intérêt	150
7.2.1	Méthodes de corrélation	150
7.2.2	Méthodes différentielles	151
7.3	Robustesse du suivi de points différentiel vis-à-vis des variations d'illumination	154
7.3.1	Méthode de suivi compensant les changements affines d'illumination .	154
7.3.2	Suivi par normalisation photométrique	155
7.4	Méthode de suivi robuste aux variations de spéularité et aux changements d'éclairage	156
7.5	Extension au suivi de motifs plus larges	157
7.6	Comparaison des méthodes de suivi vis-à-vis des variations d'illumination . . .	158
7.6.1	Résultats de simulation	159
7.6.2	Résultats d'expérimentation	165
7.7	Conclusions	178
8	L'invariance couleur	181
8.1	Modélisation de la couleur	182
8.2	Invariants dédiés aux matériaux diélectriques, inhomogènes et opaques	184
8.2.1	Illuminant blanc	184
8.2.2	Illuminant coloré	186

8.3	Invariants adaptés aux matériaux conducteurs	188
8.4	Conclusion	191
9	Suivi de points couleur robuste aux variations d'illumination	193
9.1	Suivi de points couleur	195
9.2	Suivi de points couleur basé sur une approximation photométrique locale . . .	196
9.3	Suivi couleur par utilisation d'invariants	197
9.3.1	Attributs existants	197
9.3.2	Attributs $A_1A_2A_3$	198
9.4	Suivi couleur par utilisation d'invariants et modélisation locale des variations de spécularité	198
9.5	Résultats expérimentaux	200
9.5.1	Détection de points	200
9.5.2	Comparaison des méthodes de suivi	200
9.6	Conclusions	204
	Conclusion	211
	Conclusion et perspectives	213
A	Mesure des paramètres de réflexion du jambon	III
B	Le modèle dichromatique	XI
C	Robustesse des méthodes de suivi compensant les variations d'illumination	XIII
C.1	Erreur provoquée par un pixel bruité sur la normalisation photométrique	XIII
C.2	Conditionnement de la méthode de suivi de Soatto	XIV
D	Comparaison des invariants	XVII
	Glossaire	XXI
	Bibliographie	XXV

INTRODUCTION GÉNÉRALE

Notre système visuel nous permet d'interpréter le monde 3D qui nous entoure et nous aide à agir sur cet environnement en fonction des informations perçues. Excepté dans des cas très précis d'illusions d'optique qui piègent notre faculté d'interprétation, nous comprenons de manière quasiment inconsciente l'agencement et la position des objets entre eux mais également notre propre position dans une scène. Quelles que soient les conditions d'éclairage, nous sommes capables de détecter des objets donnés, de les suivre du regard dans le temps, mais également de leur associer une sémantique inhérente aux propriétés physiques des objets (lisses, rugueux, transparents). Nous leur présumons une matière (bois, métal, plastique), nous leur associons un mouvement, une couleur. Pourtant, toutes ces informations n'arrivent pas telles quelles sur la rétine, mais résultent d'une succession de phénomènes complexes issus tout d'abord de l'interaction de la lumière avec la matière puis de la projection des stimuli lumineux sur la rétine. Ces phénomènes aboutissent finalement à des sensations colorées que le cerveau interprète.

Or la sensibilité de l'œil est très variable d'un individu à l'autre et par conséquent la couleur s'avère largement subjective. Pour compliquer les choses, notre culture et notre origine influent sur la dénomination des couleurs (rose ou fuschia, vert ou kaki). Ensuite, notre capacité d'interprétation relève bien entendu de facteurs culturels, acquis après de longues années. Finalement, le processus visuel humain et l'interprétation visuelle par le cerveau, largement étudiés par les neurobiologistes, n'ont pas encore délivré tous leurs mystères.

Forts de ce constat, nous comprenons aisément que la mise au point d'un système visuel artificiel qui permette d'atteindre le même niveau d'interprétation que le cerveau est loin d'être évidente. D'une part, l'aspect subjectif de la couleur rend difficile la définition de systèmes de mesure universels. Ensuite, certaines difficultés que le système visuel humain semble contourner à merveille peuvent être très pénalisantes en vision par ordinateur. Nous sommes capables par exemple de détecter un même objet, qu'il soit transparent ou non, quels que soient le contenu de la scène, les conditions d'éclairage, la forme de l'objet ou encore son orientation. De même, notre système visuel montre une certaine constance, c'est-à-dire une certaine adaptabilité, vis-à-vis de la perception d'un objet transparent, quelle que soit la couleur du fond sur lequel il est imprimé. À l'inverse, cette tâche peut s'avérer très ardue à mettre en œuvre en vision par ordinateur. Il en est de même lorsqu'il s'agit de suivre un élément en mouvement dont les variations d'orientation et les conditions d'éclairage modifient l'apparence au cours du temps. De manière générale, comprendre les mécanismes physiques des phénomènes d'interaction

lumière-matière-observateur tels que la transparence, l'apparition d'ombres, de spécularités, les changements d'éclairage, ne peut que faciliter la mise en œuvre d'algorithmes de traitement et d'analyse d'images, soit en modélisant ces phénomènes soit en calculant des attributs qui leur sont invariants. En outre, les attributs couleur facilitent très souvent cette tâche. Ainsi, cette thèse s'intéresse à l'apport de la couleur et des modèles de réflexion pour l'extraction et le suivi de primitives. Plus précisément, deux problématiques distinctes sont abordées, l'une concernant l'apport de la couleur pour la détection de primitives de type région par segmentation, l'autre se focalisant sur l'apport des modèles de réflexion et de la couleur pour le suivi de primitives.

D'une part, la détection d'un élément transparent dans une image peut être facilitée par l'utilisation de la couleur et par la définition d'invariants appropriés. Plus concrètement, notre problématique consiste à extraire une primitive couleur de type marquage à l'encre ou à la peinture, quelles que soient ses conditions d'impression. Bien que le système visuel humain permette une certaine constance colorimétrique, ce phénomène peut s'avérer problématique en traitement d'images dès lors qu'il s'agit de segmenter un marquage à l'encre, quelle que soit la concentration d'encre utilisée, et/ou quelle que soit la couleur du support.

D'autre part, cette thèse s'intéresse à la robustesse du suivi de primitives vis-à-vis des variations d'illumination. La plupart des techniques de suivi existantes suppose que les objets sont lambertiens, admettant ainsi que la dynamique d'une scène se résume à l'aspect géométrique et non photométrique. Quelques approches proposent une compensation d'un modèle affine des variations d'illumination dans un voisinage proche du point à suivre. À notre connaissance, leur fondement n'a pas été justifié clairement à ce jour, ce que nous proposons d'effectuer dans ce manuscrit. Dans ce cadre d'étude, nous montrons l'apport de la couleur et des modèles de réflexion à la robustesse du suivi de primitives vis-à-vis des variations d'illumination, en proposant des techniques de suivi appropriées.

Les travaux décrits dans cette thèse ont été réalisés au sein de l'unité de recherche *Technologie des équipements agro-alimentaires* du département *Écotechnologies et Agrosystèmes* du Cemagref¹. Les axes de recherche abordés dans ce département ont pour dénominateur commun le thème de la sécurité alimentaire, que ce soit au niveau de l'utilisation des agrosystèmes ou de la consommation des produits agro-alimentaires. La sécurité passe notamment par la mise au point de technologies visant à réduire certains impacts néfastes sur l'environnement, telles que les émissions de polluants. Elle passe également par la sécurité bactériologique ou sanitaire, par la qualité des produits destinés à la consommation par le biais d'analyses non invasives (par imagerie IRM par exemple), et enfin par la traçabilité des produits, qui n'est autre qu'un moyen indirect d'assurer au consommateur la qualité du produit. Ce dernier thème nous intéresse particulièrement puisque nos travaux s'intègrent dans un projet de traçabilité des produits porcins dans les abattoirs. Après des études sur l'identification fiable des pièces de porc obtenues après découpe de la carcasse, la filière porcine a préconisé le tatouage de chacune d'entre elles à l'encre alimentaire, afin de conserver trace de la provenance de l'animal. Avant l'expédition du produit hors de l'enceinte de l'abattoir, l'identifiant qui lui est associé doit être lu et enregistré. Cette tâche, qui n'est aujourd'hui que partiellement réalisée, peut être automatisée par l'utilisa-

¹Pour une description plus détaillée des activités du Cemagref, nous invitons le lecteur à consulter le site <http://www.cemagref.fr>.

tion des techniques de vision artificielle. Après que l'identifiant soit détecté par segmentation d'images, la caméra est déplacée en face de celui-ci de manière à en assurer la lecture par un algorithme de reconnaissance de caractères. D'une part, la détection de l'identifiant doit être assurée à l'aide des travaux réalisés dans cette thèse en ce qui concerne l'extraction de primitives. D'autre part, le déplacement de la caméra en face de ce marquage, déplacement contrôlé par asservissement visuel (voir par exemple [Collewet et Chaumette, 2002] ou [Alhaj, 2004]), peut souffrir des variations de spécularités à la surface de la pièce de porc. Cette tâche doit être améliorée en exploitant la technique de suivi de primitives proposée.

Ainsi, ce document propose trois contributions détaillées ci-dessous.

- Premièrement, nous étudions la transparence d'une encre colorée imprimée sur un support de couleur. Dans un premier temps, nous proposons des attributs pseudo-invariants vis-à-vis de la couleur du support, les *coefficients de concentration*, qui correspondent à l'opacité définie par le modèle de convergence de Metelli [Metelli, 1970] relatif à l'effet de transparence. Ensuite, nous définissons des attributs pseudo-invariants vis-à-vis de la quantité d'encre, les *ratios de concentration*. Ceux-ci sont exploités dans le cadre d'un algorithme de segmentation de marquages à l'encre et permettent ainsi une adaptativité des traitements vis-à-vis de la concentration d'encre et des variations de couleur de support. Finalement, l'approche proposée est validée dans le cadre de la traçabilité de la viande porcine où un identifiant doit être détecté sur des pièces de porc. À l'aide de cet exemple, nous comparons quelques méthodes de segmentation mises en œuvre et montrons la pertinence de l'utilisation d'attributs invariants.

- Ensuite, l'étude des modèles de réflexion a permis d'améliorer la robustesse du suivi de points d'intérêt vis-à-vis des variations de spécularités et des variations d'éclairage. Ainsi, à partir de l'étude des modèles de réflexion standards, nous déduisons les hypothèses sur lesquelles elles se basent. Nous proposons ensuite une méthode de suivi de points basée sur une modélisation locale des variations photométriques, celles-ci étant approximées par une fonction continue autour du point à suivre. Des expérimentations sur divers types d'objets, soumis à des changements d'illumination ou non, prouvent la robustesse de notre approche en comparaison par rapport aux méthodes existantes. De la même manière, le suivi d'objets plans peut être amélioré par une modélisation photométrique plus complète compensant les variations spatiales de changements de contraste et de spécularités.

- Nous abordons également le problème du suivi de points dans des images en couleur. En effet, on conçoit aisément que ce type d'images est susceptible de contenir plus d'information discriminante que des images de luminance. Nous améliorons la robustesse du suivi de points vis-à-vis des variations d'illumination. Ainsi, la méthode développée dans les images de luminance a été étendue au cas de la couleur. Ensuite, contrairement aux images d'intensité où les corrections d'illumination ne peuvent être que locales par le biais d'une modélisation photométrique, l'information trichromatique permet de déterminer en chaque point un pseudo-invariant photométrique et donc d'assurer une correction globale vis-à-vis des variations d'illumination. Ainsi, certains de ces attributs ont été exploités dans le cadre du suivi de points. D'autre part, nous améliorons l'utilisation des invariants les plus pertinents par une modélisation photométrique. Finalement, des expérimentations sur séquences réelles permettent de comparer les différentes approches et de prouver l'apport de l'approche proposée.

Ce document est architecturé en trois parties. La première présente quelques fondements relatifs à l'interaction entre la lumière, la matière et l'observateur, qui nous semblent nécessaires à la compréhension de la suite du manuscrit. La deuxième partie décrit les travaux relevant de l'extraction de primitives couleur, tandis que la troisième partie se consacre au suivi de primitives à partir d'images de luminance ou de couleur.

- *Partie I : L'interaction lumière-matière-observateur*

La première partie décrit les phénomènes physiques à l'origine de la formation d'une image sur le capteur. La luminance et la couleur résultent de l'interaction entre trois éléments : la lumière, la matière et l'observateur. Cette partie est composée de deux chapitres.

- Le chapitre 1 explique la formation du stimulus lumineux à partir de la lumière et des propriétés optiques du matériau. Une mesure de ces phénomènes d'interaction lumière matière peut être obtenue par le biais de la réflectance. Celle-ci peut également être modélisée simplement en fonction des propriétés optiques du matériau.

- Le chapitre 2 se focalise sur l'information disponible au niveau du capteur, à savoir la luminance et la couleur. Il aborde ainsi les concepts de photométrie et de colorimétrie. Puisque la couleur relève de la subjectivité, des normalisations ont été proposées pour aboutir à des systèmes de mesure. Il en existe un grand nombre, qui se différencient par leur propriétés et leurs techniques de calcul. Nous décrivons quelques-uns de ces espaces couleurs.

- *Partie II : Apport de la couleur pour l'extraction de primitives*

La deuxième partie s'intéresse à l'apport de la couleur pour l'extraction de primitives. Elle se scinde en trois chapitres.

- Le chapitre 3 dresse un état de l'art concernant les méthodes de segmentation d'images couleur. Il met en lumière la diversité des approches proposées, qui se basent soit sur l'extraction des contours soit sur la recherche d'homogénéités dans l'espace des couleurs ou dans la représentation géométrique de l'image. En comparaison aux images de luminance, l'utilisation de la couleur dans cette tâche offre une plus grande quantité d'information discriminante entre les différentes entités présentes dans l'image.

- Le chapitre 4 s'intéresse au phénomène de transparence de l'encre et en déduit des attributs couleur invariants : les coefficients de concentration et les rapports de coefficients de concentration. Nous proposons un algorithme de détection de marquages à l'encre par segmentation couleur, basée sur l'utilisation de ces attributs. Nous démontrons ainsi tout l'intérêt de l'apport de la couleur dans une tâche d'extraction de primitives.

- Le chapitre 5 propose la mise en œuvre de la méthode de segmentation proposée dans le cadre applicatif de contrôle de la traçabilité de produits porcins. Afin de montrer tout l'intérêt de l'approche proposée, notamment vis-à-vis de son caractère générique, nous la comparons avec une technique de segmentation développée dans l'un des espaces couleur décrits dans le chapitre 2.

- *Partie III : Apport des modèles de réflexion et de la couleur pour le suivi de primitives.*

La troisième partie du document se focalise sur le suivi de primitives dans les séquences d'images sujettes à des variations de spécularités et/ou d'éclairage. Elle se compose de quatre chapitres. Les deux premiers s'intéressent à l'utilisation de modèles d'illumination pour assurer la robustesse du suivi dans les images de luminance, tandis que les deux suivants explorent l'apport de la couleur et des invariants couleur.

- Le chapitre 6 propose de modéliser les variations d'illumination se produisant dans une séquence d'images, à partir des modèles de réflexion présentés dans le chapitre 1. De cette modélisation sont déduites des approximations photométriques valables localement dans l'image. Nous étudions finalement la validité de ces modèles en fonction de la géométrie de la scène.

- Dans le chapitre 7, le modèle d'illumination exposé dans le chapitre 6 est exploité dans une tâche de suivi de primitives de type point dans les images de luminance. Elle assure une bonne robustesse du suivi vis-à-vis des variations d'illumination. Une variante dédiée à des zones d'intérêt plus larges est également proposée.

- Le chapitre 8 consiste en un état de l'art sur les attributs couleur invariants vis-à-vis des variations d'illumination. Ils sont classés suivant les propriétés optiques des matériaux auxquels ils s'appliquent.

- Enfin, le chapitre 9 propose d'exploiter ces attributs dans le cadre du suivi de primitives points dans les images couleur. Nous proposons également l'extension à la couleur de la méthode décrite dans le chapitre 7. Puisque la plupart des invariants couleur ne compense pas parfaitement tout type de variation d'illumination, nous décrivons une approche alliant l'utilisation d'invariants couleur avec une modélisation photométrique.

Ce manuscrit se termine par une conclusion et par des perspectives. Par ailleurs, les annexes comprennent notamment un glossaire. Les termes qui y figurent sont écrits en italique et indiqués par une astérisque de la manière suivante : *mot**.

Première partie

**L'interaction
lumière-matière-observateur**

INTRODUCTION

Selon le champ d'activité considéré, que ce soit le domaine de l'art ou de l'industrie, la couleur ne correspond pas à une même réalité pratique. D'un point de vue artistique, elle est utilisée pour susciter des émotions, pour provoquer des sensations et relève donc de la subjectivité, tandis qu'au niveau industriel il s'agit d'un outil qu'il est nécessaire de maîtriser de manière objective et scientifique.

Néanmoins, la sensation de lumière ou de couleur perçue par l'œil ou par un capteur est issue de phénomènes physiques complexes qui sont eux universels. Ils résultent d'une interaction entre trois éléments : la lumière, la matière et l'observateur. Ainsi, une surface observée successivement sous deux conditions d'éclairage différentes peut engendrer deux apparences visuelles différentes. De la même manière, deux objets de même forme en aluminium et en terre cuite, que l'on soumet aux mêmes conditions d'éclairage, ne produisent pas la même image. Leur apparence visuelle est différente.

Les phénomènes physiques de l'interaction entre la lumière et la matière sont à présent bien compris. Ils peuvent être mesurés et même modélisés. Ce n'est pas tout à fait le cas de la perception qu'en a un observateur, puisqu'elle est liée à la psychologie de l'individu, et est donc difficilement quantifiable.

Bien que la vision de la couleur procure des informations non mesurables objectivement, l'essor industriel a accéléré l'étude de la colorimétrie et la création de systèmes de mesure standardisés, à savoir les espaces couleur.

L'INTERACTION LUMIÈRE-MATIÈRE

Ce chapitre introduit les concepts relatifs à la lumière, au comportement optique des matériaux et à la radiométrie. Il présente également les modèles de réflexion les plus couramment utilisés en vision par ordinateur, qui permettent de déterminer la manière dont les matériaux réfléchissent la lumière en fonction de leurs caractéristiques et des conditions d'éclairage.



L'interprétation visuelle de notre environnement relève incontestablement du domaine de la subjectivité. Il suffit d'en juger par la grande diversité des systèmes de vision animale et de leur fonctionnement. En outre, chaque individu possède une sensibilité à la lumière qui lui est propre, de même que la perception des couleurs n'est pas une faculté universelle. On sait que chez l'Homme, elle peut par exemple être altérée par une anomalie de dyschromatopsie, telle que le daltonisme par exemple.

Ce qui est vrai pour les capteurs de vision naturelle l'est également pour les capteurs de vision artificielle, qui sont plus ou moins sensibles à certaines longueurs d'onde (capteurs infrarouge, caméras couleur, caméras multispectrales). Néanmoins, dans tous les cas, l'information lumineuse reçue sur le capteur provient de phénomènes physiques complexes relatifs à l'interaction entre la lumière et la matière, qui sont quant à eux universels.

Ainsi, un grand nombre de paramètres influent directement sur la lumière réfléchi dans une direction donnée par la surface d'un objet : l'angle d'éclairage, le nombre de sources lumineuses utilisées, leur géométrie, leurs caractéristiques *photoniques**, la position du capteur par rapport à l'objet observé et aux sources lumineuses, mais également la nature physique de l'objet. En outre, une surface observée successivement sous deux éclairages différents engendrera deux stimuli différents.

Dans ce chapitre, nous nous intéressons à l'interaction lumière-matière, sans prendre en compte l'interprétation du stimulus visuel par le capteur, c'est-à-dire sans considérer l'aspect subjectif lié à la mesure. Ces phénomènes se mesurent par *radiométrie**. Dans un premier temps,

nous abordons l'aspect physique de l'interaction lumière-matière en rappelant les phénomènes lumineux, les caractéristiques optiques des matériaux et en leur décrivant un moyen de mesure : la réflectance. Nous nous focalisons ensuite sur les modèles de réflexion couramment utilisés en vision par ordinateur.

1.1 La physique de l'interaction lumière-matière

1.1.1 La lumière

En 1924, Louis de Broglie [de Broglie, 1924] découvre que la lumière repose sur une dualité onde-corpuscule. Les rayons lumineux sont des ondes électromagnétiques définies par leur fréquence ν , et leur *longueur d'onde** λ . La lumière peut être vue comme une superposition d'ondes *monochromatiques** caractérisées par leurs longueurs d'onde et par l'énergie qu'elle transportent. La lumière est produite soit en chauffant un objet jusqu'à incandescence (lampe tungstène, soleil), soit en excitant des atomes ou molécules par une décharge électrique (lampe fluorescente), et comme toute onde électro-magnétique, elle peut être *polarisée**. Sa particularité est d'être visible par une grande partie du règne animal. La lumière peut se caractériser d'un point de vue énergétique et d'un point de vue spectral.

1.1.1.1 Grandeurs énergétiques

Précisons ici quelques termes associés aux grandeurs énergétiques relatives à la lumière.

L'*énergie* E en joule d'une onde électromagnétique est proportionnelle à sa fréquence $E = h\nu$ et inversement proportionnelle à la longueur d'onde $E = hc/\lambda$ où h est la constante de Planck (6.6310^{-34}) et c la célérité de la lumière.

Le *flux énergétique* Φ_e , exprimé en watt correspond à l'*énergie* traversant une surface dans un sens donné et par unité de temps. C'est en fait l'équivalent du débit dans le cas de la propagation des fluides.

L'*éclairage énergétique* \mathcal{E} en un point d'une surface infiniment petite $d\mathcal{A}$ correspond au flux énergétique $d\Phi_e$ infiniment petit à travers cette surface, de sorte que $\mathcal{E} = d\Phi_e/d\mathcal{A}$. Il s'exprime en watt/m². Dans les ouvrages anglo-saxons, l'éclairage énergétique est appelé *irradiance*. Par abus de langage, ce terme devient également largement utilisé dans les publications francophones.

L'*intensité énergétique* I_e d'une source est le flux énergétique $d\Phi_e$ d'une source dans une direction donnée à l'intérieur d'un angle solide élémentaire $d\Omega$: $I_e = d\Phi_e/d\Omega$.

La *luminance énergétique* \mathcal{L} est le rapport de l'intensité énergétique élémentaire dI_e sur l'aire de la surface apparente vue de l'émetteur, celle-ci étant obtenue par $d\mathcal{A} \cos \theta_s$, en considérant que θ_s est l'angle entre la normale \mathbf{n} à la surface $d\mathcal{A}$ et la direction d'émission, tel que le montre la figure 1.1. Ainsi, la luminance énergétique \mathcal{L} est exprimée par :

$$\mathcal{L} = \frac{dI_e}{d\mathcal{A} \cos \theta_s} \quad (1.1)$$

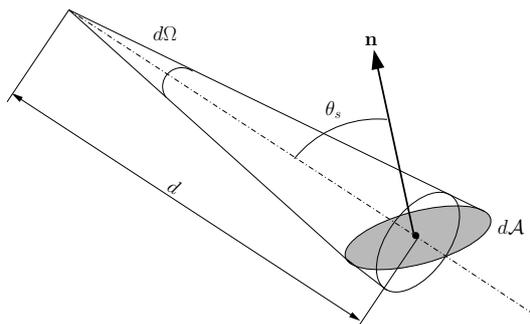


FIG. 1.1 – Illustration de la surface apparente vue de l'émetteur.

Son unité est le watt/sr/m². Dans les ouvrages anglo-saxons, elle est généralement appelée *radiance* [Desvignes, 1991].

1.1.1.2 Répartition spectrale

Chacune des sources lumineuses se distingue par sa *répartition spectrale d'énergie* $E(\lambda)$, qui est la quantité d'énergie émise par intervalle de longueur d'onde (1nm). Les grandeurs énergétiques exprimées dans la section précédente sont valides dans le domaine spectral. On parle alors d'éclairement spectral (ou irradiance spectrale), d'énergie spectrale, de flux énergétique spectral, de luminance énergétique spectrale (ou radiance spectrale).

Une source peut être quantifiée en terme de *température de couleur*, exprimée en Kelvins (K), qui représente la température à laquelle il faudrait porter un corps noir pour obtenir une répartition spectrale d'énergie identique à celle de la source en question.

L'illumination d'un objet provient généralement d'un ensemble de sources lumineuses ; elle est alors appelée *illuminant**. Certains illuminants, correspondant à des conditions d'observation courantes, ont été normalisés par la CIE¹ : on les appelle les *illuminants normalisés*. On distingue entre autre les illuminants A (lampes à incandescence 2856 K), B (soleil direct 4874K), C (jour moyen avec UV² 6504K), D (l'illuminant D_{65} notamment, est très couramment utilisé ; il représente la moyenne des lumières du jour durant une journée sans UV 6774K), E et F. La figure 1.2 montre la luminance énergétique (en mWatt/sr/m²) de quelques-uns de ces illuminants : A, B, C et D_{65} . Pour plus de détails, le lecteur pourra se référer par exemple à [Dordet, 1990, Vandenbroucke, 2000].

Les propriétés énergétiques et spectrales de la lumière étant définies, étudions à présent l'interaction de la lumière avec la matière ainsi que les principaux phénomènes lumineux.

¹Commission Internationale de l'Eclairage.

²Ultra-violet.

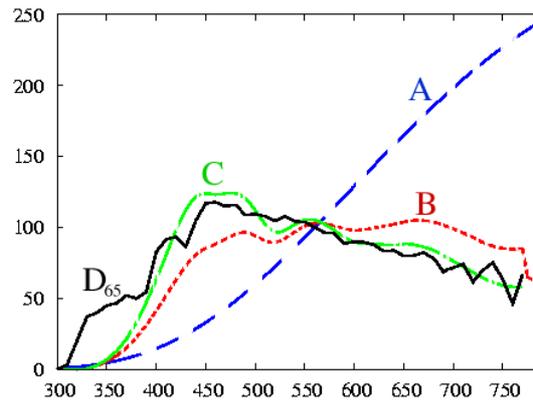


FIG. 1.2 – Luminance énergétique (en 10^{-3} watt/sr/m²) de quelques illuminants normalisés, en fonction des longueurs d'onde (en nm).

1.1.2 La matière

La lumière en provenance directe d'une source lumineuse, appelée alors *source primaire*^{*}, interagit avec une surface avant d'être réémise. De même, cette énergie réémise dans l'environnement peut interagir avec d'autres surfaces, celles-ci devenant alors des *sources lumineuses secondaires*^{*}. Les phénomènes de réflexion ne peuvent s'expliquer qu'en considérant les caractéristiques optiques intrinsèques des matériaux, puisqu'elles sont directement liées au type de surface (lisse ou rugueuse), et aux propriétés physiques des matériaux. Certaines caractéristiques optiques d'un matériau peuvent être traduites par le biais de l'*indice complexe de réfraction* et du *coefficient de Fresnel* :

- L'indice complexe de réfraction d'un milieu i est exprimé par $\hat{n}_i = n_i + i\kappa$ où n_i et κ sont deux réels traduisant respectivement le pouvoir réfractif³ et le pouvoir absorbant⁴ du matériau vis-à-vis de l'onde électro-magnétique. Ils s'expriment directement en fonction des caractéristiques physiques du matériau, c'est-à-dire des termes de *perméabilité magnétique*^{*}, de *permittivité électrique*^{*} et de *conductivité*^{*} du matériau [Brun, 2002]. Ils dépendent également de la longueur d'onde, de sorte qu'il est possible de définir un indice complexe de réfraction spectral.

- Le coefficient de Fresnel ρ_f permet de traduire la fraction de l'onde électro-magnétique qui est réfléchi à la surface du matériau.

Les matériaux peuvent être regroupés en deux grandes familles suivant leurs propriétés optiques et leurs paramètres (l'indice complexe de réfraction et le coefficient de Fresnel) : les matériaux *conducteurs* et les matériaux *diélectriques*. On se référera à la figure 1.3 pour comprendre les différents phénomènes optiques considérés.

Les **matériaux conducteurs** laissent très peu l'onde incidente les pénétrer, que ce soit par *absorption* des rayons lumineux par la matière, ou par réfraction. Ils possèdent donc un co-

³La part d'énergie de l'onde électro-magnétique qui traverse le matériau sans être atténuée.

⁴La part d'énergie de l'onde électromagnétique qui est absorbée par la matière, transformée en effet Joule.

efficient de Fresnel élevé, et un indice complexe de réfraction possédant une partie réelle n_i . L'énergie qui n'est pas réfléchi est transformée en courant de surface et produit un effet Joule. Puisque la réflexion s'opère essentiellement à la surface, par *réflexion spéculaire**, la rugosité de surface influe directement sur cette spécularité. Une surface très conductrice et très lisse est idéalement spéculaire et la loi de réflexion de Descartes s'applique à l'échelle macroscopique (voir figure 1.3a). Une partie (ou l'ensemble) de l'énergie d'un rayon lumineux d'angle d'incidence θ_i est alors réfléchi dans une seule direction, d'angle $-\theta_i$. C'est le cas des miroirs par exemple. Les matériaux plus rugueux diffusent la lumière autour de la direction de réflexion spéculaire parfaite. L'ensemble des rayons réfléchis forme alors ce que l'on appelle le lobe spéculaire (voir figure 1.3b). Le cas des métaux est tout à fait adapté pour illustrer ce phénomène, puisqu'un même métal, plus ou moins poli, réfléchira la lumière de manière plus ou moins idéalement spéculaire.

Les **matériaux diélectriques** sont des matériaux peu conducteurs. Une grande partie de l'énergie lumineuse est *absorbée** par la matière. Celle-ci est ensuite réémise par *réflexion diffuse* dans toutes les directions de l'espace (voir figure 1.3c) et/ou transmise par réfraction (voir figure 1.3d). Leur coefficient de Fresnel est faible, et leur indice de réfraction possède une partie imaginaire forte. Le phénomène lumineux dépend alors de la composition interne du matériau.

- Dans le cas de matériaux opaques, la réflexion diffuse est prépondérante. La plupart d'entre eux sont inhomogènes, c'est-à-dire qu'ils sont composés de particules possédant des caractéristiques différentes, d'indices de réfraction différents. Il s'agit généralement de divers pigments qui absorbent totalement ou partiellement la lumière à certaines longueurs d'onde, et qui sont à l'origine d'une série de réflexions en chaîne. Finalement une partie de la lumière est réémise par diffusion, et éventuellement par transmission mais de façon moins prépondérante (comme le plastique ou le bois).

- Les matériaux homogènes tels que le verre possèdent un indice de réfraction réel $\hat{n}_i = n_i$. Ainsi les phénomènes de *réfraction* et de *transmission* des rayons sont prépondérants. La réfraction est régie par les lois de réfraction de Descartes (voir figure 1.3d). Lorsqu'un rayon lumineux atteint avec un angle d'incidence θ_i une surface de séparation de deux milieux transparents, une fraction est transmise par réfraction dans le milieu, avec un angle θ_r :

$$n_1 \sin \theta_i = n_2 \sin \theta_r \quad (1.2)$$

où n_1 est l'indice de réfraction du milieu d'où est issu le rayon et n_2 celui du milieu de réfraction. Quant au phénomène de transmission, il décrit le processus de propagation du rayon lumineux à travers l'objet. Notons qu'il faut distinguer la transmission diffuse, propre aux matériaux translucides de la transmission régulière, propre aux matériaux transparents.

Selon [Hunter et Harold, 1987], 90 % des matériaux existants font partie de la classe des matériaux opaques non métalliques (papier, céramique, peinture, plastique opaque), c'est-à-dire les diélectriques inhomogènes, qui provoquent principalement une réflexion diffuse de la lumière et dans une moindre mesure une réflexion spéculaire.

Ainsi, l'interaction lumière-matière se traduit par une modification de la lumière par la matière d'un point de vue spatial, par redirection de l'onde incidente. Elle se caractérise également par une altération spectrale, puisque les pigments absorbent la lumière pour certaines plages

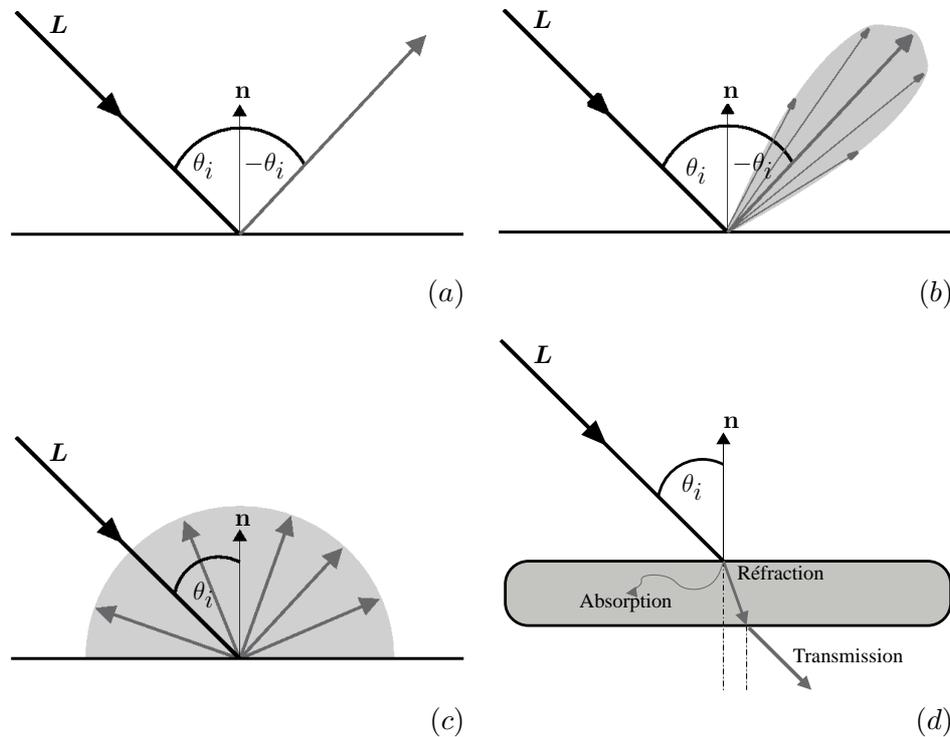


FIG. 1.3 – Le phénomène de réflexion. (a) Réflexion spéculaire idéale. (b) Réflexion spéculaire non idéale, apparition d'un lobe spéculaire. (c) Réflexion diffuse. (d) Phénomènes d'absorption, de réfraction et de transmission.

de longueurs d'ondes. Ces phénomènes complexes peuvent être traduits par une mesure de la réflectance spectrale du matériau.

1.1.3 La réflectance : une mesure de l'interaction lumière-matière

Considérons une scène dont la géométrie est décrite sur la figure 1.4. La surface dA d'aire unité est située à l'origine du repère XYZ . Soit \mathbf{n} la normale coïncidant avec le vecteur \mathbf{Z} , \mathbf{L} le vecteur de la direction de la source et \mathbf{V} la direction de l'observateur. Dans un univers 3D, deux angles sont nécessaires pour spécifier de manière unique chaque direction \mathbf{L} et \mathbf{V} . Les angles sphériques θ_i (angle zénithal) et ϕ_i (angle azimuthal) permettent de caractériser la direction du rayon incident tandis que θ_r et ϕ_r décrivent l'orientation du rayon réémis par la surface dans la direction considérée. Soit ρ l'angle entre la normale et la bissectrice \mathbf{B} entre les directions d'illumination \mathbf{L} et d'observation \mathbf{V} . La *réflectance* est définie comme le rapport de la radiance sur l'irradiance (voir le paragraphe 1.1.1.1 page 12). On appelle *réflectance spectrale* $\mathcal{R}(\lambda)$

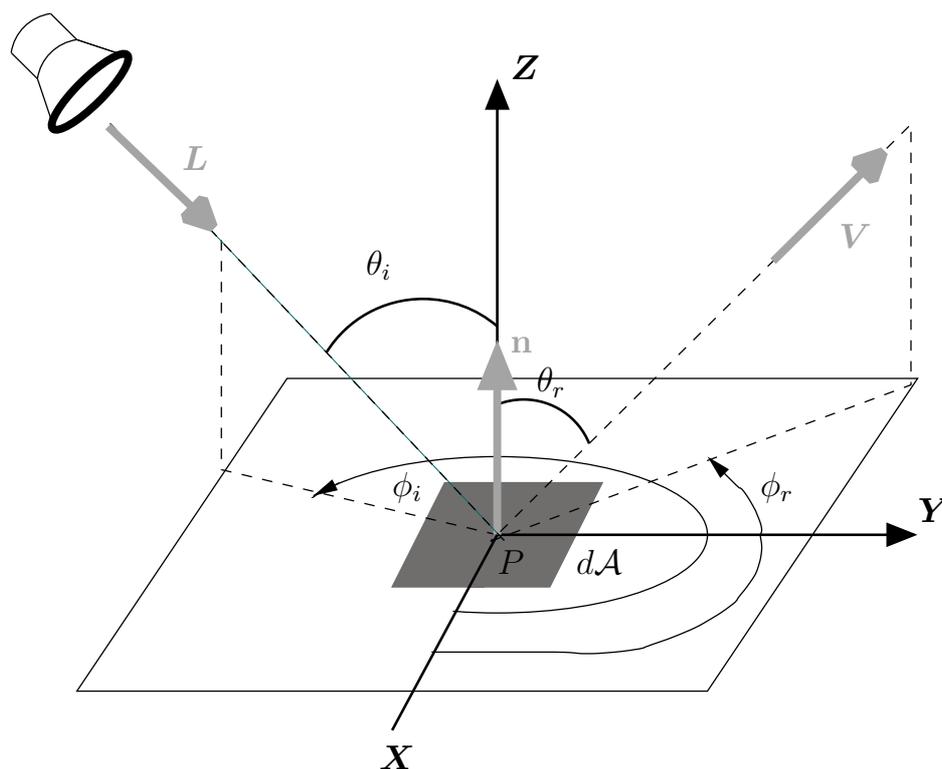


FIG. 1.4 – Paramètres géométriques décrivant le phénomène d'interaction lumière-matière-observateur.

d'une surface le rapport de la radiance spectrale $\mathcal{L}(\lambda)$ émise par un échantillon de surface dA sur l'irradiance spectrale $\mathcal{E}(\lambda)$ reçue par ce même échantillon de surface.

$$\mathcal{R}(\lambda) = \frac{\mathcal{L}(\lambda)}{\mathcal{E}(\lambda)} \quad (1.3)$$

La figure 1.5a montre un exemple de spectre de reflectance de couenne de jambon, les mesures étant exprimées en pourcentage de lumière réfléchi. L'aire sous la courbe correspond à la quantité d'énergie réfléchi, les longueurs d'ondes absorbées correspondent à des valeurs de reflectance faibles. La couenne absorbe plus de lumière dans les faibles longueurs d'onde.

Enfin, la *reflectance spectrale bidirectionnelle*, plus communément appelée BSRDF (pour Bidirectional Spectral Reflectance Distribution Function) décrit la fonction de distribution de la reflectance spectrale d'un matériau en fonction de la géométrie de la scène, c'est-à-dire de la direction d'éclairage et de la direction d'observation. Elle est déterminée par les propriétés structurales et optiques de la surface d'intérêt. Notons que l'on désigne par BRDF la reflectance directionnelle, lorsqu'elle est mesurée pour l'ensemble du spectre, et non en une longueur d'onde donnée. La BSRDF (exprimée en sr^{-1}) se définit comme le rapport de la radiance spectrale sur l'irradiance spectrale à la surface du matériau :

$$\mathcal{R}_\lambda(\theta_i, \phi_i, \theta_r, \phi_r; \lambda) = \frac{\mathcal{L}(\theta_r, \phi_r, \lambda)}{\mathcal{E}(\theta_i, \phi_i, \lambda)} \quad (1.4)$$

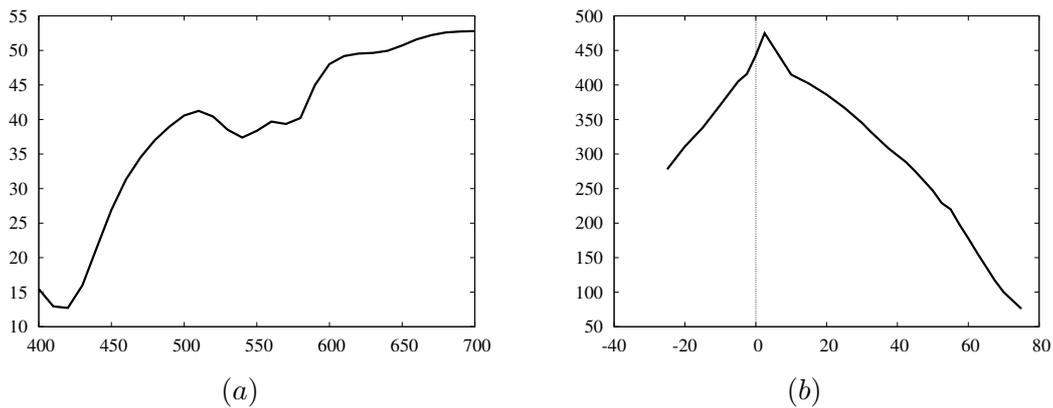


FIG. 1.5 – (a) Reflectance spectrale de la couenne ($\times 10^2$). (b) Exemple de mesure de la reflectance bidirectionnelle du jambon pour un angle d'observation de -45 par rapport à la normale à la surface ($\times 10$).

La connaissance de la BSRDF et de la radiance spectrale de l'illumination, et de la géométrie de la scène permet de connaître la radiance spectrale réfléchiée en fonction de la radiance de la source.

L'expression de la BSRDF est connue dans quelques cas idéaux. Le comportement idéal d'une surface diélectrique est d'émettre une proportion constante de l'énergie incidente, et d'absorber le reste de l'énergie. Sa BRDF ne dépend pas de la position de l'observateur, et vaut $\mathcal{R}(\theta_i, \phi_i, \theta_r, \phi_r) = k_l/\pi$, où k_l est l'*albedo** du matériau. Dans le cas d'une surface spéculaire idéale, il y a réflexion totale des rayons incidents, tel que le stipulent la loi de réflexion de Descartes. Dans ce cas, la BRDF est également connue (voir [Hussenet, 2002] par exemple). Cependant, la plupart des diélectriques réfléchissent une part de la lumière par réflexion spéculaire, ce qui amène généralement à considérer des formulations de BRDF très complexes. Ainsi, bien que quelques travaux aient permis de calculer la BRDF de nombreux matériaux, il est souvent nécessaire de procéder à une mesure de celle-ci. La figure 1.5b montre un exemple de mesure de reflectance bidirectionnelle de la couenne d'un jambon, pour une direction d'observation $\theta_r = -45$. On trouvera dans l'annexe A les explications concernant la mise en œuvre de cette mesure.

Comme nous allons le voir à présent, les phénomènes de réflexion peuvent également être modélisés en fonction de la géométrie de la scène et des paramètres caractéristiques de l'état de surface du matériau, à savoir les paramètres de rugosité.

1.2 Modélisation des phénomènes d'interaction lumière-matière

Il s'agit, en connaissant la forme d'un objet, sa texture, sa rugosité et la position des sources lumineuses, de modéliser la manière dont l'objet réfléchit la lumière dans une certaine direction (celle de l'observateur). D'autre part, la compréhension des mécanismes de réflexion lumineuse permet d'extraire des propriétés de surface, notamment de rugosité, à partir de l'intensité de la lumière réfléchiée par l'objet. Étant donnée la prédominance des objets diélectriques inhomo-

gènes, les principaux modèles se rapportent à ce cas de figure. Ils négligent pour la plupart les phénomènes de réfraction et de transmission et prennent surtout en compte les phénomènes de diffusion et de réflexion spéculaire. Or, comme nous l'avons précisé dans la section 1.1.2, le phénomène de réflexion spéculaire dépend en partie de l'état de surface du matériau. Il convient donc de prendre en compte dans un premier temps la rugosité de surface des matériaux.

1.2.1 Les modèles de rugosité de surface

La description de la rugosité d'une surface passe par la modélisation de ses micro-aspérités. Il existe principalement deux approches classiques de modélisation des irrégularités de surface : par une distribution normale des hauteurs d'aspérités ou par des micro-surfaces. Nous abordons également le facteur de rugosité, introduit par Nayar.

1.2.1.1 Modélisation par une distribution normale des hauteurs des aspérités

Soient h la hauteur des irrégularités de surface et σ_h l'écart-type de leur distribution (voir figure 1.6). La rugosité peut être modélisée par la densité de probabilité $\rho_h(h)$ de h telle que :

$$\rho_h(h) = \frac{1}{\sqrt{2\pi}\sigma_h} e^{-\frac{h^2}{2\sigma_h^2}} \quad (1.5)$$

L'écart-type σ_h traduit finalement la rugosité de la surface. Plus l'écart-type est faible moins il y a occurrence d'aspérités importantes. Cependant, ce modèle ne permet pas de modéliser la distance entre deux pics. De ce fait, Beckmann *et al.* [Beckmann et Spizzichino, 1987] y apportent quelques améliorations, en introduisant un paramètre traduisant la corrélation $\mathcal{C}(d)$ entre les altitudes de deux points de surface séparés par une distance d . Il s'écrit :

$$\mathcal{C}(d) = e^{-\frac{d^2}{\tau^2}} \quad (1.6)$$

où τ représente la distance de corrélation pour laquelle $\mathcal{C}(d) = e^{-1}$.

1.2.1.2 Modélisation par micro-facettes

Dans le modèle de Torrance-Sparrow [Torrance et Sparrow, 1967], la surface rugueuse est approximée par un ensemble de micro-surfaces planes et lisses de mêmes dimensions, qui sont décrites par les angles β_s entre leurs normales \mathbf{n}_s et la normale à la macro surface \mathbf{n} (voir figure 1.7). Supposons que la surface a_s associée à une micro-surface soit nettement supérieure à la longueur d'onde $\lambda \ll a_s$. Dans ces conditions, l'aspect électromagnétique de la lumière peut être négligé, et la rugosité peut être modélisée en appliquant les lois de l'optique géométrique sur chacune des micro-facettes. La répartition des orientations associées aux micro-surfaces est modélisée par une fonction de distribution normale de moyenne nulle et d'écart-type ς :

$$\rho_\beta(\beta) = c e^{-\frac{\beta^2}{2\varsigma^2}} \quad \text{avec} \quad c = \left(\int_0^{\frac{\pi}{2}} e^{-\frac{\beta^2}{2\varsigma^2}} d\beta \right)^{-1} \quad (1.7)$$

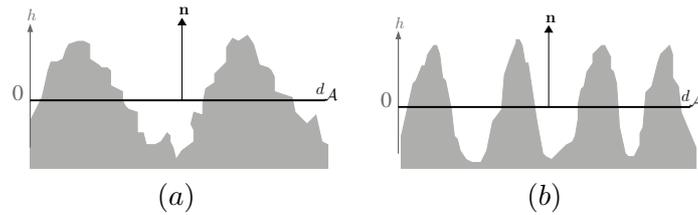


FIG. 1.6 – Modélisation par une distribution normale des hauteurs des aspérités. (a) Grande distance de corrélation. (b) Faible distance de corrélation.

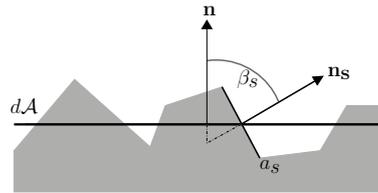


FIG. 1.7 – Modélisation par micro-facettes. La normale à la macrosurface et la normale à la microsurface forment un angle β_s .

De manière à simplifier la représentation de la rugosité, la distribution des micro-surfaces est supposée unidimensionnelle. Cela implique que le modèle soit invariant par rotation autour de \mathbf{n} au point considéré et par conséquent, que la distribution des normales \mathbf{n}_s soit identique quelle que soit la rotation autour de \mathbf{n} . L'écart-type des angles β_s , noté ici ζ , correspond indirectement à la rugosité. En effet, plus la surface est rugueuse et plus les variations au niveau de l'orientation β_s des micro-facettes sont importantes.

1.2.1.3 Le facteur de rugosité de Nayar

En se basant sur la modélisation de la rugosité par distribution des hauteurs de rugosité, Nayar [Nayar *et al.*, 1991] introduit un terme de rugosité g exprimé de la manière suivante :

$$g = \left(2\pi \frac{\sigma_h}{\lambda} (\cos(\theta_i) + \cos(\theta_r)) \right)^2 \quad (1.8)$$

σ_h n'est autre que l'écart-type de la distribution des hauteurs de rugosité exprimé dans la section 1.2.1.1. Plus cet écart-type est important vis-à-vis de la longueur d'onde λ et plus la surface est rugueuse. Ainsi, le niveau de rugosité se déduit du facteur g [Hussenet, 2002] :

- si $g \ll 1$, la surface est très lisse ;
- si $g \simeq 1$, la surface est moyennement rugueuse ;
- si $g \gg 1$, la surface est très rugueuse.

Ainsi, les modèles de rugosité déterminent la manière dont est réfléchi la lumière, et constituent la base des modèles de réflexion existants.

1.2.2 Les modèles de réflexion

Les lois de Descartes ne s'appliquant qu'à des surfaces idéales, un nombre significatif de travaux a porté sur la modélisation plus réaliste des phénomènes de réflexion. Elle s'appuie sur plusieurs paramètres relatifs à l'objet, tels que la rugosité, l'*albedo**, ainsi qu'à la géométrie de la scène (directions d'éclairage et d'observation).

Nous présentons ici les modèles de réflexion les plus couramment utilisés en vision par ordinateur. Précisons que leurs formalismes ne prennent pas directement en compte les inter-réflexions entre objets, les ombres projetées et les masquages.

1.2.2.1 Le modèle lambertien

Le modèle lambertien [Lambert, 1760] est le modèle le plus simple pour décrire les phénomènes de réflexion puisqu'il suppose l'absence totale de spécularités. Il permet de relier la radiance en un point P dans une direction d'observation donnée à l'angle d'incidence $\theta_i(P)$, aux caractéristiques du matériau (*via* l'albedo ou reflectance diffuse $\mathcal{R}_b(\lambda, P)$) et au spectre d'illumination $\mathcal{E}(\lambda, P)$ au point P :

$$\mathcal{L}^L(\lambda, P) = \begin{cases} \mathcal{E}(\lambda, P)\mathcal{R}_b(\lambda, P) \cos \theta_i(P) & \text{si } \theta_i(P) \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ 0 & \text{sinon} \end{cases} \quad (1.9)$$

Remarquons que cette relation ne prend pas en compte la direction de l'observateur θ_r , le spectre de lumière émise ne varie donc pas selon les angles de vue. Le modèle lambertien n'est pertinent que si une partie suffisamment significative de l'onde incidente pénètre dans le matériau avant d'être retransmise par réflexion diffuse. Il est donc mieux adapté aux matériaux diélectriques inhomogènes. Cette modélisation est une simplification du comportement de la reflectance, mais sa linéarité en fait une modélisation intéressante. Notons qu'il n'est valable que pour des angles d'incidence $|\theta_i| < \frac{\pi}{2}$.

1.2.2.2 Le modèle de Phong

Les améliorations apportées par Phong [Phong, 1975] sont heuristiques, et ne se basent sur aucune hypothèse physique. Néanmoins, sa simplicité en fait un modèle facilement exploitable et largement exploité en synthèse d'image ou en analyse d'image. La radiance dans une direction donnée est la suivante :

$$\mathcal{L}^P(\lambda, P) = \mathcal{E}(\lambda, P)\mathcal{R}_b(\lambda, P) \cos \theta_i(P) + K_s(\lambda, P) \cos^n(\rho(P)) + K_a(\lambda, P) \quad (1.10)$$

À l'instar du modèle lambertien, le modèle de Phong comprend un terme correspondant à la réflexion diffuse. Il introduit de surcroît un terme décrivant le lobe de réflexion spéculaire, comme l'illustre la figure 1.8a. La valeur de ce terme dépend à la fois de la rugosité du matériau n et de l'angle ρ entre \mathbf{B} et \mathbf{n} . K_s et K_a dépendent du spectre de l'illuminant. Ce modèle fait l'hypothèse d'un éclairage ponctuel.

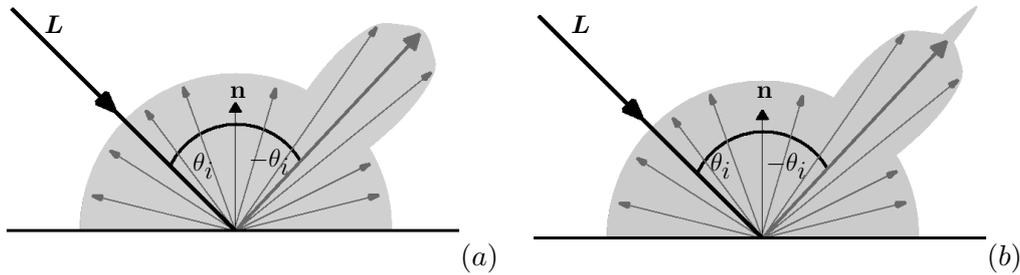


FIG. 1.8 – (a) Réflexion mixte, diffuse et spéculaire. (b) Illustration du modèle de Nayar : présence de réflexion diffuse, d'un lobe spéculaire et d'un pic spéculaire.

1.2.2.3 Le modèle de Torrance-Sparrow

À la différence des modèles précédents, le modèle de réflexion proposé par [Torrance et Sparrow, 1967] se base sur l'optique géométrique. Puisqu'il néglige l'aspect électromagnétique de la lumière, sa validité est restreinte aux cas où la taille des aspérités de surface excède largement la longueur d'onde lumineuse. Ces irrégularités de surface sont modélisées par des micro-surfaces, comme expliqué dans le paragraphe 1.2.1.2. La radiance au point P est exprimée comme suit :

$$\mathcal{L}^T(\lambda, P) = \begin{cases} \mathcal{E}(\lambda, P)\mathcal{R}_b(\lambda, P) \cos \theta_i(P) + \frac{K_s(\lambda, P)}{\cos(\theta_r(P))} e^{(-\rho^2(P)/2\zeta^2)} & \text{si } \theta_i(P) \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ 0 & \text{sinon} \end{cases} \quad (1.11)$$

où ζ correspond au paramètre de rugosité du modèle des micro-facettes. Contrairement au modèle lambertien, le modèle de Torrance-Sparrow permet de décrire la réflexion de matériaux conducteurs, par l'addition du second terme, associé à la composante spéculaire de la réflexion.

1.2.2.4 Le modèle de Beckmann-Spizzichino

Le modèle de Beckmann-Spizzichino [Beckmann et Spizzichino, 1987] décrit la réflexion produite à la surface du matériau en utilisant la théorie des ondes électromagnétiques, ce qui engendre des relations complexes que nous ne décrivons pas ici. Retenons toutefois que les aspérités de surface suivent une distribution normale, comme le décrit le paragraphe 1.2.1.1 page 19, et que ce formalisme se base sur d'autres hypothèses que celle de la rugosité, puisqu'il suppose notamment que l'onde incidente est plane et polarisée perpendiculairement.

Puisqu'il décrit les réflexions à la surface du matériau, le modèle de Beckmann-Spizzichino est mieux adapté aux matériaux conducteurs. Il est valide aussi bien pour des surfaces lisses que pour des surfaces rugueuses. Ce modèle est constitué de deux termes, l'un proportionnel à e^{-g} qui correspond au pic de réflexion spéculaire (qui apparaît uniquement dans le cas de surfaces extrêmement lisses) tandis qu'un second terme est associé au lobe spéculaire. Afin d'obtenir

des précisions supplémentaires sur ce modèle, le lecteur pourra se référer par exemple à [Brun, 2002].

1.2.2.5 Le modèle de Nayar

Le modèle de Nayar [Nayar *et al.*, 1991] s'avère intéressant dans la mesure où il réunit trois composantes issues des différents modèles abordés précédemment, à savoir le modèle de Torrance-Sparrow et le modèle de Beckmann-Spizzichino en ce qui concerne la description des composantes spéculaires, et le modèle Lambertien en ce qui concerne la composante diffuse :

$$\mathcal{L}^N(\lambda, P) = a\mathcal{L}^T(\lambda, P) + b\mathcal{L}^B(\lambda, P) + c\mathcal{L}^L(\lambda, P) \quad \text{avec } a, b, c \text{ réels} \quad (1.12)$$

La réflexion est donc décrite par des composantes diffuse et spéculaire ainsi qu'un pic spéculaire, comme l'illustre la figure 1.8b. Il opère quelques simplifications sur le modèle de Torrance-Sparrow, en stipulant dans un premier temps, que le coefficient K_s de l'équation (1.11) est constant ; dans un deuxième temps il limite le protocole expérimental à un système où l'observateur est fixe, seule la source étant mobile. Dans ce cas les angles θ_r et ϕ_r sont constants. Le pic spéculaire de Beckmann-Spizzichino peut être approximé par une fonction de Dirac, égale à 1, dans la direction spéculaire et 0 ailleurs.

En prenant en compte ces approximations, le modèle de Nayar s'exprime de la manière suivante :

$$\mathcal{L}^N(\lambda, P) = \frac{K_s(\lambda, P)}{\cos(\theta_r(P))} e^{-\frac{\alpha^2(P)}{2\zeta^2(P)}} + K_{ss}(\lambda, P) \delta(\theta_i(P) - \theta_r(P)) \delta(\theta_r(P)) + K_d(\lambda, P) \cos(\theta_i(P)) \quad (1.13)$$

Les premier et second termes contribuent à la modélisation de la réflexion spéculaire. Ils sont basés respectivement sur les modèles de Torrance-Sparrow et de Beckmann-Spizzichino. Le dernier terme est relatif à la composante lambertienne. Ainsi, le modèle de Nayar est valide pour une plus grande variété de matériaux, qu'ils soient conducteurs ou non. Ceci justifie son appellation de *modèle unifié*.

La section suivante présente le *modèle dichromatique*, qui exprime la réflexion de manière qualitative, à l'aide de termes géométriques et de termes spectraux.

1.2.2.6 Une synthèse qualitative : le modèle de réflexion dichromatique

La théorie de Kubelka-Munk permet l'analyse des phénomènes de diffusion de la lumière à une échelle macroscopique du matériau. Cette théorie explique la formation du spectre réfléchi d'un corps coloré, en se basant sur les fonctions de diffusion et d'absorption propres au matériau, en admettant que la lumière est diffusée de manière isotropique à l'intérieur du matériau. Le modèle de réflexion dichromatique, proposé par Shafer [Shafer, 1985], se base sur cette théorie, à savoir que tout diélectrique inhomogène, uniformément coloré et opaque, réfléchit la lumière soit sous forme de radiance de surface (*interface reflection* en anglais) soit sous forme de radiance interne (*body reflection* en anglais). Dans le premier cas, le rayon réfléchi conserve plus ou moins les caractéristiques spectrales de la lumière incidente, et de ce fait le stimulus de

couleur correspond à la couleur de la source. Quant à la réflexion interne, elle résulte de la pénétration des rayons dans l'objet puis de sa diffusion par les particules (pigments) du matériau de l'objet. Le stimulus de couleur correspond dans ce cas à la couleur de l'objet.

Le modèle de réflexion dichromatique, permet ainsi de dissocier ces deux phénomènes, de telle sorte que la radiance d'un objet $\mathcal{L}(\lambda, P)$ devient la somme de la radiance issue de l'interface $\mathcal{L}_s(\lambda, P)$ et de la radiance interne $\mathcal{L}_b(\lambda, P)$:

$$\mathcal{L}(\lambda, P) = \mathcal{L}_s(\lambda, P) + \mathcal{L}_b(\lambda, P) \quad (1.14)$$

Chacune des radiances \mathcal{L}_r et \mathcal{L}_s est constituée du produit de deux termes, tel que :

$$\mathcal{L}(\lambda, P) = l_s(\lambda, P)m_s(P) + l_b(\lambda, P)m_b(P) \quad (1.15)$$

Les termes m_b et m_s sont des facteurs d'échelle dépendant uniquement de la géométrie de la scène. $l_b(\lambda, P)$ correspond à la radiance spectrale associée au phénomène de réflexion diffuse. Enfin, $l_s(\lambda, P)$ est couramment associé au spectre de l'illuminant : $l_s(\lambda, P) = \mathcal{E}(\lambda, P)$. [Healey, 1989] en propose une formulation généralisée :

$$\mathcal{L}(\lambda, P) = \begin{cases} m_s(P)l_s(\lambda, P) & \text{pour les matériaux conducteurs} \\ m_s(P)l_s(\lambda, P) + m_b(P)l_b(\lambda, P) & \text{pour les diélectriques inhomogènes opaques} \end{cases} \quad (1.16)$$

Cette généralisation se base sur l'hypothèse stipulant que les métaux de conductivité élevée laissent très peu l'onde incidente les pénétrer, et possèdent donc une composante lambertienne négligeable par rapport à la composante spéculaire. Le modèle de réflexion dichromatique ne s'applique qu'aux scènes éclairées par une seule source de lumière, sans aucune lumière ambiante et sans interréflexion. Ce modèle paraît très restreint mais s'applique en revanche à de nombreux matériaux.

Nous verrons dans le chapitre 8 tout l'intérêt de ce modèle dichromatique dans certaines applications de vision par ordinateur.

Avec l'essor du domaine de la synthèse d'images, maintes autres modélisations de la réflectance ou de la radiance ont vu le jour, mais ces formalismes ne sont pas aisément exploitables en vision par ordinateur. Le lecteur intéressé par ce sujet d'étude pourra se référer à d'autres bibliographies, telle que celle présentée dans [Claustres, 2003].

1.3 Conclusion

La lumière est une onde électromagnétique dont la plage de longueurs d'onde coïncide avec le domaine de sensibilité de l'œil humain. L'interaction d'une onde électromagnétique, telle que la lumière, avec la matière se manifeste à la fois par une déviation des rayons incidents, une modification du spectre et/ou une absorption d'énergie. Nous avons vu comment les caractéristiques de ces phénomènes s'expliquent par la constitution interne du matériau ainsi que par son aspect de surface.

Par le biais d'une mesure de la réflectance d'un matériau, il est possible de caractériser son comportement optique. Bien que les phénomènes de réflexion considérés soient parfois

complexes, ils sont modélisables de différentes manières. Alors que certains de ces modèles sont empiriques (le modèle de Phong, ou le modèle dichromatique par exemple), d'autres se basent sur les propriétés électromagnétiques de la lumière (Beckmann-Spizzichino) ou sur des propriétés d'optique géométrique (Torrance-Sparrow).

Alors que ce premier chapitre a permis de comprendre les phénomènes à l'origine de l'irradiance de la surface dans une direction d'observation et d'aborder les modèles de réflexion les plus largement utilisés en vision par ordinateur, le chapitre 2 porte sur les informations issues du capteur : la luminance et la couleur.

LA LUMINANCE ET LA COULEUR

Alors que le chapitre 1 a présenté la formation d'un stimulus lumineux à partir des phénomènes de réflexion de la lumière sur les matériaux, ce chapitre permet de comprendre la formation de la couleur et de la luminance. Il aborde ainsi les concepts de photométrie et de trichromie. Nous verrons ensuite qu'un grand nombre d'espaces de représentation de la couleur a été défini.



Dans le chapitre précédent, l'aspect subjectif du phénomène visuel, c'est-à-dire la façon dont le stimulus lumineux visible est perçu par un observateur, n'a pas été pris en compte. Il s'est focalisé sur les aspects radiométriques, sans aborder la notion de couleur ou encore la luminance réellement perçue par le capteur ou l'observateur. Or à notre sens, avant d'exploiter les images couleur, il nous semble essentiel de comprendre comment la couleur se définit et comment elle est représentée. Notons qu'il existe des capteurs multispectraux et des capteurs définis pour des plages de longueurs d'onde non visibles (infra-rouge par exemple), que nous ne traitons pas dans ce manuscrit.

La couleur a longtemps été un domaine réservé aux artistes, de part la sensation totalement subjective qu'elle représente. Au début du XX^e siècle, le seul moyen qu'avaient les industriels d'effectuer les suivis colorimétriques était d'employer des personnes spécialisées dans les mélanges de couleurs. Depuis cette période, les recherches en colorimétrie ont élaboré une définition objective de la couleur, en se basant sur la composition spectrale de la lumière.

Dans un premier temps, nous décrivons la luminance en montrant le lien entre la radiance de la scène, définie dans le chapitre 1, et l'irradiance à la surface du capteur. Ensuite, nous nous intéressons à la couleur et à l'effet trichrome de la sensation visuelle. Enfin, quelques espaces couleur sont décrits. Définis initialement dans le but de mesurer la couleur, ces espaces sont à présent largement exploités dans le cadre du traitement d'image.

2.1 La luminance

Contrairement à la *radiométrie*, qui décrit d'un point de vue physique l'énergie émise par une source, la *photométrie* exprime d'un point de vue psychophysique la luminance perçue par un capteur, celui-ci pouvant être artificiel (un capteur CCD par exemple) ou naturel (l'œil humain par exemple). Dans le premier cas, le lien entre les données radiométriques et les données photométriques a été démontré. Dans le deuxième cas, la luminance perçue par l'œil humain s'avère subjective, ce qui a nécessité la définition d'un *observateur standard*.

2.1.1 Luminance mesurée par un capteur CCD

Le processus de propagation de la lumière dans un objectif de distance focale f et d'ouverture d est décrit par la figure 2.1. Les rayons lumineux issus d'une surface dA de l'objet convergent avec un angle γ vers la lentille de l'objectif et se projettent sur une surface S du capteur. La relation entre la radiance de l'objet observé \mathcal{L} et l'irradiance reçue sur le capteur, que nous notons \mathcal{I}_c , a été obtenue par Horn [Horn, 1986] :

$$\mathcal{I}_c(\lambda) = K_c \mathcal{L}(\lambda) \quad (2.1)$$

K_c est un paramètre dépendant de la géométrie de l'objectif. Il s'écrit :

$$K_c = \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \gamma \quad (2.2)$$

Notons que K_c ne dépend aucunement de la longueur d'onde. Dans le cas où le capteur est considéré suffisamment petit par rapport à la scène observée, on considère que $\gamma \simeq 0$ et donc $\cos(\gamma) = 1$. Par conséquent, à focale f et ouverture d fixes, la valeur K_c est une constante.

Soit un point de la scène P se projetant sur le capteur en un point p . Notons qu'il existe différents modèles de projection (voir la thèse [Lingrand, 1999] par exemple). La luminance lumineuse $f(p)$ mesurée par le capteur au point p dépend de la sensibilité de celui-ci. Elle est obtenue par l'intégration suivante :

$$f(p) = K_c \int_{\lambda_{\min}}^{\lambda_{\max}} \mathcal{S}(\lambda) \mathcal{L}(\lambda, P) d\lambda \quad (2.3)$$

Ainsi, une image f est définie comme l'ensemble structuré regroupant les luminances mesurées en chaque point p du capteur.

2.1.2 Luminance perçue par l'œil humain

Étant donné que la luminance perçue par un observateur dépend de chaque individu, il a été nécessaire de définir un observateur standard. De manière générale, on peut cependant affirmer que l'œil humain est sensible aux radiations électromagnétiques de longueurs d'onde comprises entre 380 et 780 nm.

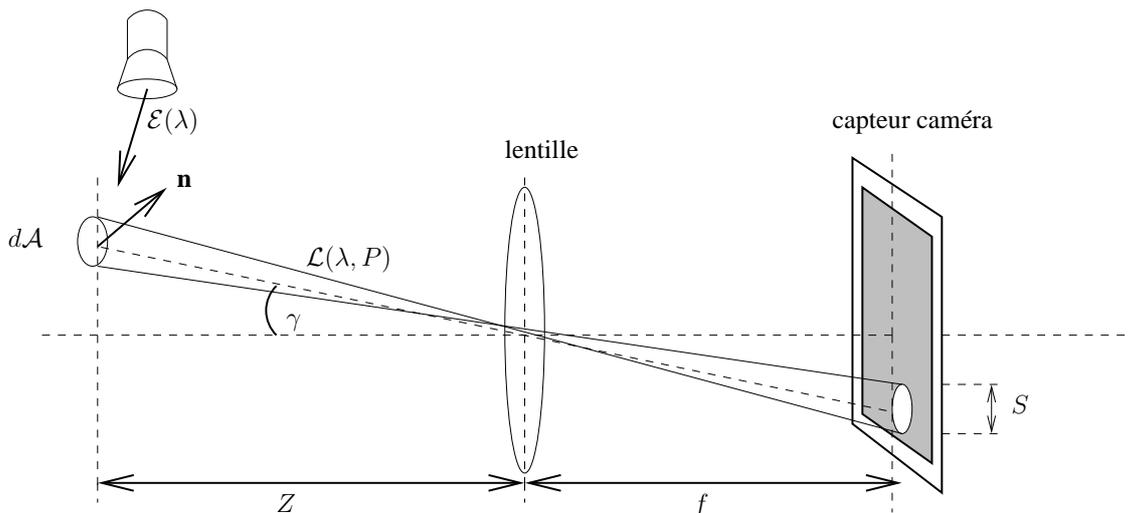


FIG. 2.1 – Réflexion d'une onde électromagnétique sur le capteur de la caméra.

L'observateur standard. En 1924, la CIE définit, à partir de plusieurs expériences psychosensorielles la *fonction d'efficacité lumineuse relative spectrale* $V(\lambda)$ décrite par la figure 2.2 (données issues de [Trouvé, 1991]), qui représente la sensibilité lumineuse de l'œil à différentes longueurs d'onde du domaine du visible en *vision photopique**. Il est ainsi démontré que l'œil ne présente pas la même sensibilité pour toutes les longueurs d'onde de cet intervalle, mais atteint un maximum pour une longueur d'onde de 560 nm . Notons qu'il existe d'autres « modes » de vision (la *vision scotopique** et la *vision mésopique**). Cet observateur standard est d'une incontestable utilité, puisqu'il permet de passer de données radiométriques à des données photométriques. Les grandeurs lumineuses sont obtenues en pondérant les grandeurs énergétiques (exposées dans le paragraphe 1.1.1.1 page 12) par la fonction d'efficacité $\mathcal{V}(\lambda)$. Notamment, la luminance lumineuse I est obtenue par pondération du stimulus $\mathcal{L}(\lambda)$ par la fonction d'efficacité $\mathcal{V}(\lambda)$:

$$I = \int_{\lambda_{\min}}^{\lambda_{\max}} \mathcal{V}(\lambda) \mathcal{L}(\lambda) d\lambda \quad (2.4)$$

Dans la suite du manuscrit, on désignera la *luminance lumineuse** par les termes intensité ou encore luminance. De la même manière, les termes *énergie lumineuse*, *éclairage lumineux** et *intensité lumineuse** désignent les grandeurs photométriques définies sur le même principe, à partir des données radiométriques de la section 1.1.1.1. L'unité de base des mesures photométriques n'est plus le watt mais le *candela**.

2.2 La couleur

Alors que la luminance perçue par l'œil dépend uniquement de l'énergie du rayon lumineux, la couleur apparente est induite par la répartition spectrale de son énergie. Dans cet intervalle de longueurs d'onde associé à la *lumière visible*, on distingue différents domaines correspondant

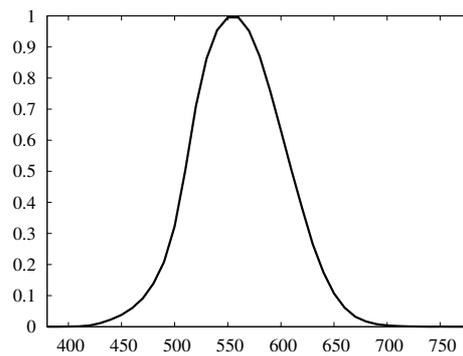


FIG. 2.2 – Fonction d'efficacité lumineuse relative spectrale $V(\lambda)$.

à des stimulations colorées distinctes. Dans l'ordre croissant de longueur d'onde, on trouve le violet, le bleu, le vert, le jaune, l'orange et le rouge.

La couleur est une sensation humaine dont l'étude psychologique et scientifique a débuté il y a plus d'un siècle. Les travaux en vision par ordinateur sont intimement liés à ces études, par la création de caméras dont la sensibilité est comparable à celle de l'observateur standard.

La trichromie. La perception humaine de la couleur est caractérisée par son aspect tridimensionnel. L'expérience montre que seules trois couleurs, appelées *primaires*, permettent de reproduire la quasi-totalité des couleurs. C'est ce qu'on appelle la *trichromie*. Les trois primaires peuvent être arbitrairement choisies sous la seule réserve qu'aucune d'elles ne puisse être reproduite par un mélange approprié des deux autres.

Helmholtz [Helmholtz, 1866] a été le premier à différencier deux types de mélanges de couleurs : le mélange par *synthèse additive**, associé au mélange de faisceaux de couleur et le mélange par *synthèse soustractive**, associé au mélange d'encre ou de peintures. Dans le premier cas, les primaires sont le rouge, le vert et le bleu (voir figure 2.3a) ; dans le second, ce sont le jaune, le cyan et le magenta (voir figure 2.3b). C'est à Maxwell [Maxwell, 1860] que l'on doit d'avoir véritablement établi les principales lois des mélanges de couleur. Du fait de cette trichromie, les couleurs peuvent être représentées dans un espace tridimensionnel, chacun des trois vecteurs de la base désignant une couleur primaire. Le vecteur couleur C est défini par une combinaison linéaire des vecteurs de la base (C_1, C_2, C_3) :

$$C = \alpha C_1 + \beta C_2 + \gamma C_3 \quad (2.5)$$

Ainsi, il est possible de déterminer expérimentalement la quantité des trois primaires C_1 , C_2 et C_3 ou *composantes trichromatiques* nécessaires à l'*égalisation** de chacun des rayonnements monochromatiques C du visible.

En se basant sur ce même principe, la CIE effectue en 1931 une synthèse des résultats expérimentaux sur la vision¹ et normalise un système permettant d'interpréter avec précision les

¹Expérimentations équivalentes effectuées par Wright en 1929 sur 10 observateurs et Guild en 1931 sur 7 observateurs.

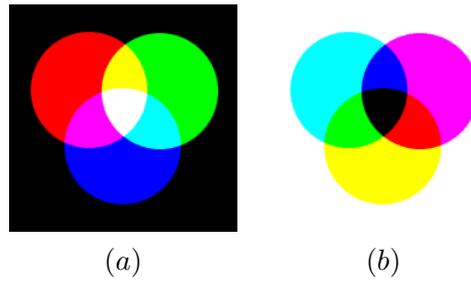


FIG. 2.3 – La trichromie. (a) Synthèse additive. (b) Synthèse soustractive.

sensations colorées. Elle se base sur trois primaires monochromatiques : le rouge ($\lambda=700$ nm), le vert (546,1 nm) et le bleu (435,8 nm). À l'issue de ces expériences ont pu être définies des *composantes trichromatiques spectrales* ou fonctions trichromatiques $\bar{r}(\lambda)$, $\bar{v}(\lambda)$ et $\bar{b}(\lambda)$, qui indiquent les quantités de chacune des primaires qui seraient nécessaires à l'égalisation d'un stimulus monochromatique de longueur d'onde λ . Ces courbes, sont données par la figure 2.4a. Elles caractérisent l'*observateur standard CIE 1931*, et se vérifient pour environ 90 % des observateurs. L'expression de l'*efficacité lumineuse relative* (voir la relation 2.4) est obtenue à partir de ces fonctions, de la manière suivante :

$$\mathcal{V}(\lambda) = 1.000 \bar{r}(\lambda) + 4.5907 \bar{g}(\lambda) + 0.0601 \bar{b}(\lambda) \quad (2.6)$$

Notons que la courbe $\bar{r}(\lambda)$ comprend une partie négative, ce qui va à l'encontre de la théorie de synthèse additive, puisqu'il est impossible d'ajouter une lumière négative. Toutes les couleurs pures monochromatiques ne peuvent donc pas être reproduites dans le diagramme chromatique *RGB*. En pratique, pour parvenir à l'égalisation, il a été nécessaire d'ajouter une quantité de rayonnement rouge au stimulus à égaliser C . Dans le seul but de pallier ce défaut, les fonctions de l'observateur standard CIE 1931 sont remplacées plus tard par les fonctions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ et $\bar{z}(\lambda)$ (voir figure 2.4b) qui sont quant à elles théoriques, puisqu'elles ne sont pas attestées par l'expérience. Elles s'obtiennent par simple transformation linéaire à partir de $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ et $\bar{b}(\lambda)$:

$$\begin{cases} \bar{x}(\lambda) = 2.7688 \bar{r}(\lambda) + 1.7517 \bar{g}(\lambda) + 1.1301 \bar{b}(\lambda) \\ \bar{y}(\lambda) = \mathcal{V}(\lambda) \\ \bar{z}(\lambda) = 0.0001 \bar{r}(\lambda) + 0.0565 \bar{g}(\lambda) + 5.5942 \bar{b}(\lambda) \end{cases} \quad (2.7)$$

L'interprétation de la couleur. Notre perception de la couleur dépend à fois du signal de couleur mais également de la manière dont ce signal est interprété par le couple œil-cerveau. En effet, l'interprétation des couleurs dépend à la fois de mécanismes neurophysiologiques mais également d'aspects psychologiques liés à l'apprentissage des couleurs qui est différent d'un individu à l'autre. La façon dont on nomme couramment les couleurs (jaune citron, vert pomme, rouge clair, fuschia) dépend à la fois de l'éducation et de la culture, et il s'avérerait difficile de faire admettre de manière universelle, et non approximative, une appellation différente pour chacune des milliers de couleurs discernables. Par contre, comme l'a introduit Maxwell, la couleur peut être décrite de manière subjective, intuitive à l'aide de trois attributs : la luminance

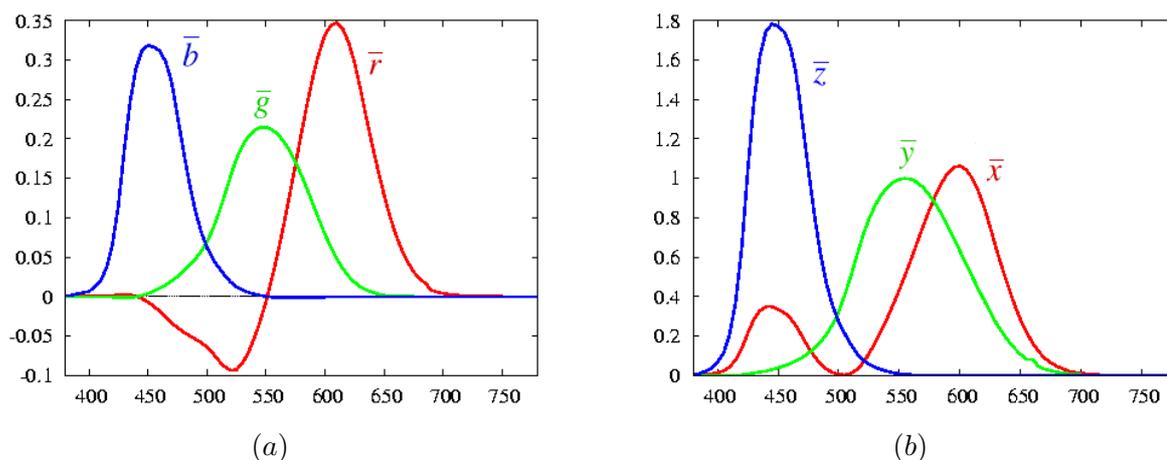


FIG. 2.4 – (a) Composantes trichromatiques spectrales $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$. (b) Composantes trichromatiques spectrales $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$.

lumineuse, dite *luminosité** ou *clarté** (une surface paraît lumineuse ou claire si elle semble émettre beaucoup de lumière, si elle est claire), la *teinte** (associée couramment à la notion de couleur : rouge, vert, jaune...) et la *saturation** ou *pureté** (on dit communément qu'une couleur est pure ou vive lorsqu'elle est très saturée, et délavée dans le cas contraire). La saturation et la teinte sont appelées *composantes de chrominance* de la couleur et reflètent l'aspect qualitatif de celle-ci, la luminosité correspondant quant à elle à l'aspect quantitatif.

Du fait d'un besoin croissant de mesure de la couleur, notamment dans les milieux industriels, de nombreux systèmes colorimétriques ont été développés. Tous ces espaces couleur découlent de la trichromie ou des attributs d'interprétation de la couleur (teinte, saturation, luminosité). La section qui suit décrit quelques-uns de ces systèmes. Précisons que nous n'avons pas cherché à effectuer une description exhaustive des espaces colorimétriques et nous avons d'ailleurs préféré ne pas aborder les systèmes colorimétriques utilisés uniquement en télévision et en synthèse soustractive (imprimerie offset, imprimantes). Pour obtenir de plus amples informations quant à ces espaces, le lecteur pourra se référer à une étude très détaillée présente dans la référence [Vandenbroucke, 2000].

2.3 La représentation de la couleur

Les espaces couleur peuvent être classés suivant quatre grandes familles : les systèmes de primaires, les systèmes perceptuellement uniformes, les systèmes perceptuels, qui sont basés sur les attributs de perception de la couleur, et enfin les systèmes d'axes indépendants. Définis à l'origine dans le cadre strict de la colorimétrie, ils sont à présent largement exploités en vision par ordinateur. Dans cette section, nous soulignons également le problème du choix de l'espace couleur en fonction du type d'application.

2.3.1 Les systèmes de primaires

2.3.1.1 L'espace RGB de la CIE

Le système RGB (pour Red Green Blue), ou RVB en français, résulte des travaux de la CIE précédemment exposés dans la section 2.2. Nous n'abordons ici que l'espace RGB de la CIE 1934, mais il faut néanmoins noter qu'il existe de nombreux espaces RGB qui se différencient entre eux par l'illuminant utilisé mais aussi par le nombre de bits de codage que requiert l'application (espaces YIQ , YUV et YC_1C_2). Ces systèmes sont utilisés par exemple en télévision.

Ainsi, les coordonnées trichromatiques RGB sont définies à partir de la distribution spectrale du stimulus lumineux $\mathcal{L}(\lambda)$ de longueur d'onde comprise entre λ_{min} et λ_{max} et des *fonctions colorimétriques spectrales* abordées dans la section 2.2 (voir figure 2.4b) :

$$R = \int_{\lambda_{min}}^{\lambda_{max}} \mathcal{R}(\lambda)\mathcal{E}(\lambda)\bar{r}(\lambda)d\lambda \quad G = \int_{\lambda_{min}}^{\lambda_{max}} \mathcal{R}(\lambda)\mathcal{E}(\lambda)\bar{g}(\lambda)d\lambda \quad B = \int_{\lambda_{min}}^{\lambda_{max}} \mathcal{R}(\lambda)\mathcal{E}(\lambda)\bar{b}(\lambda)d\lambda \quad (2.8)$$

Ces équations, qui simulent la perception trichrome de l'observateur standard, nous permettent de faire le parallèle avec les capteurs couleur, qui possèdent également leur propre sensibilité spectrale. Ainsi, une caméra tri-CCD comprend trois capteurs, R , G et B de sensibilités respectives $S_R(\lambda)$, $S_G(\lambda)$ et $S_B(\lambda)$. Les coordonnées RGB sont obtenues de la même manière que dans (2.8), en remplaçant toutefois les fonctions colorimétriques spectrales par les fonctions de sensibilité des capteurs. Généralement, ces dernières correspondent approximativement aux fonctions colorimétriques spectrales, mais des différences peuvent apparaître suivant les constructeurs. Soulignons le fait que, d'après (2.8), deux combinaisons différentes des variables \mathcal{E} , \bar{r} , \bar{g} et \bar{b} peuvent aboutir à une valeur identique de coordonnées RGB . C'est ce que l'on appelle le *métamérisme**.

Toute couleur s'exprime dans l'espace RGB par un vecteur C de coordonnées (R, G, B) tel que :

$$C = RR^* + GG^* + BB^* \quad (2.9)$$

où R^* , G^* et B^* correspondent à la base associée au système RGB . Chaque sommet du cube (voir figure 2.5) correspond à une couleur caractéristique : le point origine O est associé au noir, le point de coordonnées $(1, 1, 1)$ au blanc et les six autres sommets correspondent aux primaires de la colorimétrie additive et de la colorimétrie soustractive. La droite passant par les points noir et blanc est l'*axe achromatique*, axe reliant tous les points associés à des couleurs issues d'un mélange parfaitement équitable des trois primaires, autrement dit des nuances de gris.

Dans cet espace couleur, la luminance L est définie par la somme des composantes trichromatiques $L = R + G + B$. Ceci induit que deux stimuli de couleur, de composantes trichromatiques R , G et B différentes, peuvent présenter la même luminance.

Maxwell a présenté un espace couleur sous la forme d'un triangle dont les sommets sont le rouge, le vert et le bleu. On retrouve ce triangle dans l'espace RGB , perpendiculaire à l'axe achromatique et dont les sommets correspondent aux extrémités des axes chromatiques (voir figure 2.5). Dans ce triangle chaque couleur est référencée par ses *coordonnées trichromatiques*,

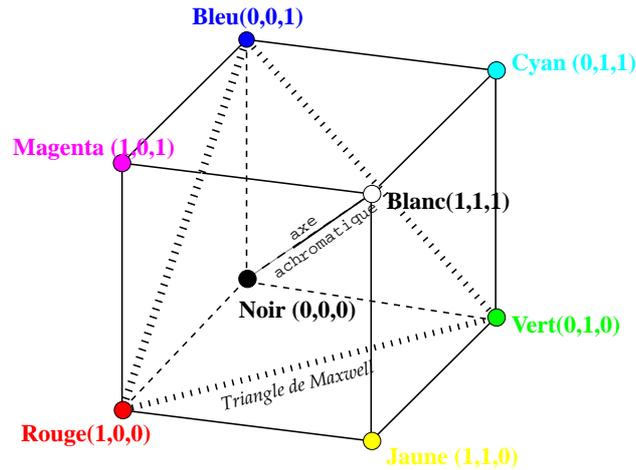


FIG. 2.5 – Représentation de la couleur dans l'espace RGB.

qui correspondent aux composantes trichromatiques normalisées par rapport à la luminance (ou coordonnées réduites) :

$$r = \frac{R}{R + G + B} \quad g = \frac{G}{R + G + B} \quad b = \frac{B}{R + G + B} \quad (2.10)$$

Puisque l'on a $r + g + b = 1$, deux coordonnées chromatiques sont suffisantes pour localiser une couleur.

Malgré l'utilité du système RGB pour les acquisitions d'images, il montre certains inconvénients, le plus gênant d'entre eux étant lié au défaut d'égalisation évoqué dans la section 2.2. Toutes les couleurs ne peuvent effectivement pas être créées à l'aide des composantes RGB . C'est le cas des pourpres par exemple, dont la synthèse trichrome introduit des valeurs de composantes négatives. À ce désavantage se rajoute celui que nous avons abordé précédemment, à savoir que la luminance est dépendante des trois composantes RGB .

Ainsi, d'autres systèmes colorimétriques ont été établis de manière à proposer une meilleure représentation de la couleur. Abordons tout d'abord le cas du système XYZ qui, tout comme le système RGB , est devenu un système de référence standardisé par la CIE.

2.3.1.2 Système de référence colorimétrique CIE 1931 XYZ et système xyY

Afin d'élaborer le nouveau système de référence XYZ , les fonctions colorimétriques $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ ont été remplacées par $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ (voir la section 2.2 ainsi que la figure 2.4b). De cette façon, les stimuli XYZ ne correspondent aucunement à un rayonnement ou à des sensations réelles, mais résultent d'une simple transformation linéaire du système RGB . Les lois simples de mélange des couleurs de l'espace RGB sont conservées. Les deux principaux défauts de l'espace RGB sont corrigés, puisque d'une part les composantes de référence XYZ permettent d'exprimer toutes les couleurs à l'aide de composantes trichromatiques positives et que, d'autre part, la primaire Y porte l'information de luminance. Effectivement, la fonction

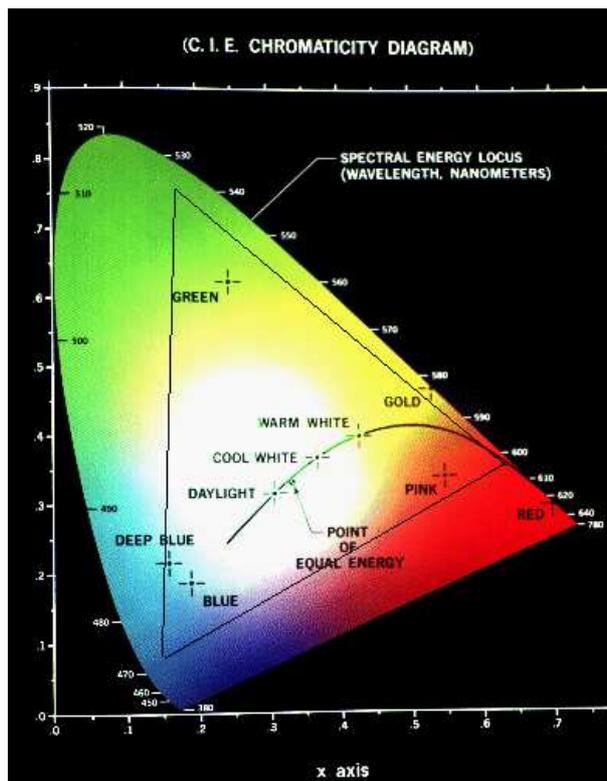


FIG. 2.6 – Diagramme de chromaticité de la CIE.

colorimétrique $y(\lambda)$ correspond à la fonction d'efficacité spectrale relative de l'œil $\mathcal{V}(\lambda)$ (voir l'équation (2.6)).

Dans le triangle des couleurs de Maxwell, situé dans le plan de luminosité constante, une couleur est identifiée par ses coordonnées trichromatiques données ci-dessous :

$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad z = \frac{Z}{X + Y + Z} \quad (2.11)$$

Elles permettent de représenter tous les stimuli de couleur dans un *diagramme de chromaticité* (voir figure 2.6), seule la chrominance des couleurs y étant représentée. Signalons également que le diagramme de chromaticité montre un intérêt pratique, que ce soit au niveau du mélange des couleurs² ou au niveau de la recherche de couleurs complémentaires³. Les « colorimétriciens » s'en servent avec succès depuis plus d'un demi-siècle. Enfin, grâce à la loi de Planck, on a accès au *locus* des coordonnées chromatiques associées à l'émission d'un corps noir, qui est visible sur la figure 2.6. Notamment, les coordonnées des illuminants standards sont bien identifiées.

Afin de définir complètement une couleur, il est nécessaire de déterminer le paramètre d'intensité or, comme nous l'avons précédemment signalé, Y correspond exactement à la sensibilité

²La couleur résultante d'un mélange est située sur la ligne reliant les deux couleurs participant au mélange.

³Il suffit de tracer une ligne entre la couleur considérée et le blanc de référence jusqu'à l'autre côté de la courbe des couleurs pures.

visuelle de l'œil humain $\mathcal{V}(\lambda)$. Un nouveau système est ainsi déterminé à partir des trois coordonnées x, y, Y . La chromaticité de la couleur est obtenue par convention à l'aide des deux premières composantes et l'intensité par la coordonnée Y , qui est portée sur un axe perpendiculaire au plan de chromaticité. La CIE adopte en 1964 un second système appelé système de référence colorimétrique supplémentaire. Il s'agit d'une amélioration du système précédent, dont les mesures expérimentales se basaient sur une région très limitée et centrale de la rétine (la *fovéa**) en utilisant un champ visuel de 2 degrés. Ainsi le nouvel espace s'appuie sur des mesures effectuées à l'aide d'un champ visuel élargi (un angle solide de 10 degrés, d'où le nom du système X_{10}, Y_{10}, Z_{10}). Les travaux récents montrent que la mesure idéale devrait reposer sur un champ visuel de 7 d'angle solide.

Les systèmes primaires présentés dans cette section comportent tous un défaut d'uniformité, c'est à dire qu'il n'y a pas de linéarité entre la distance entre deux couleurs, mesurée dans le système, et l'*écart chromatique** perçu réellement par l'observateur standard. En d'autres termes, une même distance couleur mesurée à deux endroits différents du système couleur peut conduire à deux écarts perceptuels différents. Ainsi, la création d'un espace uniforme fut requise par beaucoup d'industriels désireux d'aboutir à une métrique objective dans l'espace des couleurs et à une automatisation des mesures d'écarts colorimétriques. C'est ainsi que de plus amples travaux ont permis d'aboutir à des systèmes perceptuellement uniformes. Ils sont décrits ci-après.

2.3.2 Les systèmes perceptuellement uniformes

Le phénomène de non-uniformité des espaces couleur est très bien mis en exergue par l'expérience de MacAdam en 1942 [MacAdam, 1942] (ellipses de MacAdam), qui est d'ailleurs à l'origine de nombreuses études sur les non-linéarités des espaces colorimétriques. En effet, il considère plusieurs points du diagramme de chromaticité (figure 2.6) et construit le *locus* des points correspondant aux couleurs perceptuellement similaires à la couleur considérée. La limite du lieu correspond donc au plus petit écart de couleur discernable par le système visuel humain. Le résultat est très probant puisque les lieux correspondent approximativement tous à des ellipses, mais des ellipses de tailles et d'orientations différentes selon l'endroit du point considéré. Ainsi, notre œil possède une plus grande sensibilité dans le vert (la courbe de sensibilité de la figure 2.2 montre un maximum en 560 nm) et permet de distinguer dans ces longueurs d'onde de faibles variations de teinte. Inversement notre œil a une plus faible sensibilité dans le bleu et le rouge.

Par la suite, la CIE a normalisé des espaces ($L^*u^*v^*$, $L^*a^*b^*$), issus d'une transformation non-linéaire des systèmes de primaires. Cependant, déterminer un système parfaitement uniforme est loin d'être simple étant donnée la nature non euclidienne de l'espace XYZ . La solution de ce problème n'a pas été trouvée à ce jour, bien que plusieurs espaces, déterminés de manière empirique ou issus de transformations de diagrammes de chromaticité, sont admis comme uniformes. Dans la majeure partie des cas, du fait de la non-linéarité de la transformation, les propriétés de mélanges de couleur sont perdues⁴.

⁴Un mélange de deux couleurs ne se trouve plus sur le segment qui les relie.

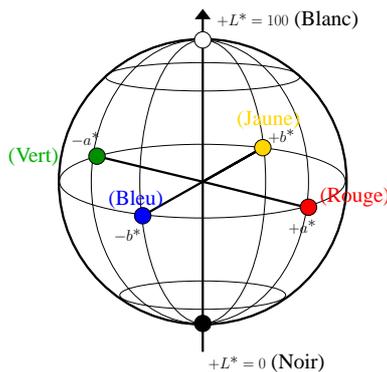


FIG. 2.7 – Systèmes uniformes. Exemple de l'espace CIE 1976 (L^* , a^* , b^*).

2.3.2.1 L'espace CIE 1964 (u^* , v^* , w^*)

Un premier espace dit uniforme, (u^* , v^* , w^*), a été mis au point par la CIE en se référant au diagramme de MacAdam. Les composantes u^* et v^* portent l'information de chromaticité, et une troisième composante, appelée *clarté psychométrique* w^* , permet de faire intervenir la luminance. Cet espace permet de définir un véritable écart colorimétrique, qui n'est autre que la distance euclidienne. Malheureusement il s'avère que ces écarts ne sont pas uniformes suivant la luminance.

Ce système n'eut pas le succès escompté car il n'apportait pas une amélioration notable aux résultats souhaités par les milieux industriels. Ainsi, il fallut attendre 1973 pour que deux espaces uniformes créés par la CIE soient jugés satisfaisants, une version modifiée du système (u^* , v^* , w^*) de 1964 : (L^* , u^* , v^*), ainsi qu'une version issue des travaux de Adams-Nickerson : (L^* , a^* , b^*), que nous décrivons à présent.

2.3.2.2 L'espace CIE 1976 (L^* , a^* , b^*)

Ce modèle, représenté sur la figure 2.7, est le standard de référence, d'où parfois l'appellation de *vraie couleur* lorsqu'une couleur est exprimée en valeurs $L^*a^*b^*$. La mise en œuvre de cet espace trouve son origine dans les travaux d'Adams et Nickerson, l'espace ANLAB 40. Les composantes (L^* , a^* , b^*) sont obtenues à partir des composantes (X , Y , Z), de la manière suivante :

$$\begin{cases} L^* & = 116(Y/Y_n)^{1/3} - 16 \text{ pour } \frac{Y}{Y_n} > 0.008856 \\ \text{ou : } L^* & = 903.3(Y/Y_n)^{1/3} \text{ pour } \frac{Y}{Y_n} \leq 0.008856 \\ a^* & = 500[f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})] \\ b^* & = 200[f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})] \end{cases} \quad (2.12)$$

où (X_n , Y_n , Z_n) sont les coordonnées du blanc de référence. La fonction f est donnée par :

$$f(x) = \begin{cases} \sqrt[3]{x} & \text{si } x > 0.008856 \\ 7.787x + \frac{16}{116} & \text{si } x \leq 0.008856 \end{cases} \quad (2.13)$$

La luminance est exprimée en pourcentage, de 0% (noir) à 100% (blanc). a^* est l'axe vert-rouge et b^* l'axe bleu-jaune.

Le système (L^*, a^*, b^*) n'est considéré qu'approximativement uniforme. Ainsi la distance colorimétrique (initialement définie par la distance euclidienne) a été sujette à des améliorations jusqu'en 1994. À présent, elle se définit à partir du *chroma*^{*}, de la saturation et de la teinte (voir [Trémeau *et al.*, 2004] par exemple).

Les systèmes uniformes font partie d'une grande famille d'espaces de luminance-chrominance. La section qui suit s'intéresse plus particulièrement à d'autres espaces de luminance-chrominance, plus proches de la perception humaine.

2.3.3 Systèmes perceptuels

Le peintre et professeur d'art Albert H. Munsell étudia les couleurs au début du vingtième siècle, et décida de les ordonner selon une méthode systématique. Pour ce faire, il utilisa les attributs associés à la perception subjective de la couleur qu'avait auparavant défini Maxwell, à savoir la *teinte* (hue en anglais), la *luminosité* (value) et la *saturation* (chroma). Munsell classa les couleurs dans un réseau cylindrique selon ces trois critères. Ce système est à la base d'un atlas encore très utilisé à l'heure actuelle, dit *Atlas de Munsell* ou *Munsell Book of Color*.

En suivant le même principe d'expression psychovisuelle des couleurs, Ostwald définit en 1916 un *Alphabet des couleurs*, celles-ci étant classées en fonction de trois variables : la *proportion de couleur*, la *proportion de blanc* et la *proportion de noir*. Il introduisit aussi l'expression particulière de *couleur pleine*, qu'il définit comme une couleur qui ne laisse percevoir qu'un seul ton (le *Hue* de Munsell) et qui ne contient aucun mélange de blanc ou de noir.

Beaucoup de systèmes perceptuels, plus compréhensibles par l'Homme, ont été créés. Certains d'entre eux se définissent en exprimant en coordonnées polaires ou cylindriques des composantes issues de systèmes luminance-chrominance. En l'occurrence, les espaces $L^*H^*C^*$ et $L^*H^*S^*$ sont respectivement issus d'une transformation des espaces $(L^*a^*b^*)$ et $(L^*u^*v^*)$. Leur représentation est donnée par la figure 2.8a, où C , H et L correspondent respectivement au chroma, à la teinte et à la luminance.

Les systèmes « humains » de perception de la couleur sont quant à eux issus d'une transformation non-linéaire du système de primaires RGB . Il existe un grand nombre d'espaces perceptuels se différenciant par les transformations nécessaires à leur obtention. Citons notamment le modèle cylindrique ISH ⁵ et le modèle triangulaire HSL ⁶. Nous avons choisi de ne développer qu'un des espaces les plus courants, le système HSV , que nous utilisons ultérieurement dans ce manuscrit.

Le modèle de cône hexagonal HSV (pour Hue-Saturation-Value) est représenté sur la figure 2.8b. La composante de teinte est matérialisée par un angle qui peut varier conventionnellement de 0 à 360°. Notons que c'est également le cas du système HLS . Les primaires additives et soustractives se situent sur les sommets des hexagones. La saturation S exprime l'éloignement

⁵Pour Intensity Saturation Hue en anglais.

⁶Pour Hue Saturation Luminance en anglais.

de la couleur vis-à-vis de l'axe achromatique, tandis que la luminosité est donnée par la composante V , valant zéro pour le noir, et atteignant la valeur maximale au maximum de clarté que peut atteindre la couleur.

$$\begin{cases} H = \arctan \left(\frac{\sqrt{3}(G - B)}{2R - G - B} \right) \\ V = \max(R, G, B) \\ S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \text{ si } \max(R, G, B) \neq 0 \text{ sinon } S = 0 \end{cases} \quad (2.14)$$

Notons que dans le cas du système HLS , la teinte est définie de la même manière, mais la saturation s'écrit :

$$S = 1 - 3 \frac{\min(R, G, B)}{R + G + B} \quad (2.15)$$

Les systèmes perceptuels montrent plusieurs avantages [Sangwine et Horne, 1998], dont le principal est d'assurer une bonne compatibilité avec la perception. D'autre part, la séparabilité des valeurs chromatiques par rapport aux valeurs achromatiques représente un grand intérêt. Notamment, utiliser une composante de chrominance, telle que la teinte par exemple, permet d'améliorer les algorithmes de traitement d'images en s'affranchissant en partie des variations d'illumination.

Néanmoins, ce type de système comporte également son lot d'inconvénients. Premièrement l'uniformité assurée dans l'espace L^*, a^*, b^* (ainsi que l'espace L^*, u^*, v^*) ne l'est plus pour les espaces de luminance-chrominance. D'autre part, ils comportent certaines singularités. Notamment, la teinte n'est pas définie pour les valeurs non saturées.

2.4 Systèmes d'axes indépendants : système de Ohta

Ohta [Ohta *et al.*, 1980] a déterminé des composantes couleur indépendantes. Pour ce faire, il utilise un échantillon représentatif de huit images et effectue une transformation de Karhunen-Loeve (KLT), qui s'avère être une référence en ce qui concerne la décorrélation des composantes RGB . Il parvient à obtenir un système d'axes qui s'avère être une bonne approximation de la transformation KLT, et dont les composantes sont décrites comme suit :

$$I_1 = \frac{R+G+B}{3} \quad I_2 = \frac{R-B}{2} \quad I_3 = \frac{2G-R-B}{4} \quad (2.16)$$

En particulier, cette simple transformation linéaire de l'espace RGB permet de s'affranchir des temps prohibitifs de la transformation KLT. Remarquons que la première composante I_1 correspond à la luminance, ce qui permet de classer ce système parmi les espaces de luminance-chrominance.

Nous avons décrit quelques espaces couleur parmi la multitude de systèmes existants. Ils possèdent tous des caractéristiques bien particulières, avantageuses ou défavorables. Devant une telle importance de choix, il s'avère bien difficile de choisir le système colorimétrique « idéal ». C'est pourquoi nous consacrons la section qui suit à cette question.

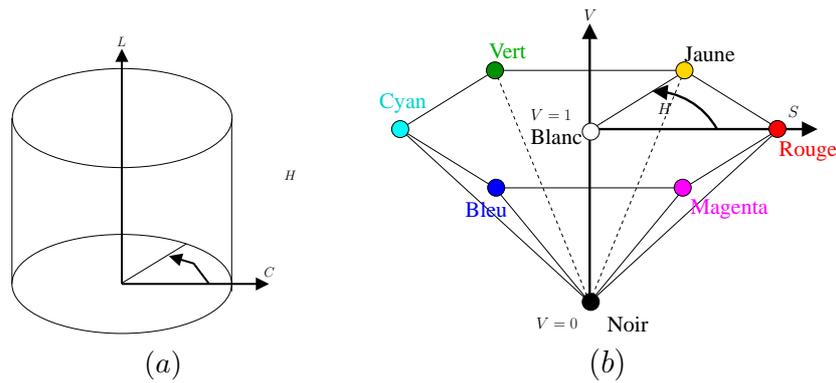


FIG. 2.8 – (a) Représentation cylindrique d'un espace perceptuel. (b) Système hexagonal HSV.

2.4.1 Choix du système couleur

Pour la plupart des études consacrées au choix du système colorimétrique dans différents cadres d'application (traitement d'images numériques [Sharma et Trussell, 1942, Vandenbroucke, 2000], invariants couleur [Berry, 1987], quantification [Brun et Tremeau, 2002]) la conclusion générale montre que le choix d'un système couleur dans le cadre d'un traitement donné reste un problème largement ouvert. Même si quelques auteurs privilégient tel ou tel espace pour effectuer un type de traitement donné, il est aisé de trouver la contradiction dans un autre ouvrage. En effet, le choix d'un espace dépend à la fois de l'application et du type d'images à traiter. De plus, la comparaison des résultats de traitement d'images obtenus à l'aide de différents espaces n'est pas une tâche aisée. On trouvera dans certains ouvrages, tel que dans [Trémeau *et al.*, 2004] une analyse détaillée des avantages et inconvénients de plusieurs espaces couleur, selon l'application visée. L'espace RGB possède des composantes directement visualisables, et reste un espace privilégié pour certaines applications temps-réel. Les espaces $(L^*a^*b^*)$ et $(L^*u^*v^*)$ sont à favoriser dès lors qu'il s'agit de mesurer des écarts de couleur, puisqu'ils permettent une meilleure relation entre distance couleur et écart colorimétrique. Néanmoins, pour être utilisés à bon escient, ils nécessitent la connaissance des conditions d'acquisition. Certains traitements nécessitent une décorrélation des composantes couleur. Dans ce cas, on privilégie généralement l'espace $I_1I_2I_3$ de Ohta. Enfin, les systèmes de type Luminance Teinte Saturation sont intéressants pour définir des invariants couleur, étant données les propriétés d'invariance de la teinte vis-à-vis des ombres portées. Par contre, leur principal défaut est la présence de singularités pour des valeurs de saturations faibles.

Finalement, la formation d'*espaces hybrides* peut constituer une solution intéressante pour le traitement d'images. Les composantes d'un espace hybride sont des attributs sélectionnés spécifiquement de manière à individualiser au mieux les différentes classes présentes dans une image donnée, au sens d'une connaissance *a priori*. Ainsi, l'espace hybride est complètement dépendant de l'image à traiter et de l'application en amont, généralement une classification. Les méthodes de détermination des composantes de couleur discriminantes se basent sur l'étude d'un ensemble de classes d'apprentissage, à partir duquel sont recherchés les attributs les plus discriminants vis-à-vis des classes, parmi un panel d'attributs prédéfinis. Une composante s'ob-

tient soit par recherche de nouvelles variables, par exemple à l'aide d'une analyse factorielle discriminante, soit par tri dans un ensemble d'attributs existants.

Dans le premier cas, il s'agit de rechercher à partir d'un ensemble d'individus d'apprentissage (les couleurs d'une image d'apprentissage) la combinaison linéaire des composantes couleur permettant de rendre compte du mieux possible de la partition initiale des classes. Les différentes variables retenues sont alors pondérées les unes par rapport aux autres.

Dans le deuxième cas, il s'agit de sélectionner un certain nombre de composantes parmi plusieurs espaces existants. Dans [Vandenbroucke, 2000], l'auteur utilise la sélection itérative. Dans un premier temps, un espace monodimensionnel composé de la composante la plus discriminante vis-à-vis des classes en présence, est créé. Ensuite, celle-ci est associée à toutes les composantes restantes pour obtenir l'espace bidimensionnel le plus discriminant, et ainsi de suite, jusqu'à obtenir un espace de dimension D fixée.

2.5 Conclusion

Contrairement aux phénomènes physiques issus de l'interaction entre la lumière et la matière, qui ont été abordés dans le premier chapitre, la mesure de la luminance et de la couleur fait intervenir la sensibilité du capteur.

Tandis que la photométrie traduit d'un point de vue psychophysique la luminance reçue par un capteur, la colorimétrie s'intéresse à l'étude de la sensation colorée ainsi qu'à sa mesure objective, par l'utilisation des propriétés de trichromie. La quantité d'espaces couleur existants est si importante, du fait de la diversité des méthodes d'analyse couleur, qu'il paraît vain de vouloir en dresser une liste exhaustive. Néanmoins, nous en avons décrit quelques-uns, suffisamment représentatifs des différents types d'espaces rencontrés dans la littérature. On distingue notamment les espaces de primaires, les espaces uniformes, les espaces perceptuels, d'axes indépendants. Pour la plupart de ces systèmes, leur création a été motivée par des applications.

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

CONCLUSION

Comme nous l'avons vu dans cette partie, les informations de luminance et de couleur dépendent tout d'abord de phénomènes physiques liés à l'interaction entre l'onde lumineuse et la matière. Ce sont à la fois les caractéristiques spectrales de l'illuminant et les propriétés optiques des matériaux qui déterminent le stimulus lumineux. Ces phénomènes sont à présent bien connus, et la mesure de la réflectance d'un matériau permet de caractériser son comportement optique. Ils peuvent d'ailleurs être modélisés de différentes manières, de façon empirique (par le biais du modèle de Phong par exemple), en se basant sur l'optique géométrique (le modèle de Torrance-Sparrow) ou sur l'électromagnétisme (Beckmann-Spizzichino).

Ensuite, la notion de couleur correspond à une information psychovisuelle qu'il est difficile de quantifier de manière tout à fait objective, puisqu'elle dépend de la sensibilité du capteur, qu'il soit naturel (le système visuel humain) ou artificiel (une caméra tri-CCD). Quelques espaces couleur, en adéquation correcte avec la sensibilité humaine, ont cependant été normalisés et s'avèrent largement utilisés en traitement d'images. Nous en avons décrit les plus pertinents.

L'objectif de l'analyse et du traitement d'images en vision par ordinateur consiste à extraire de l'information pertinente quant à la structure d'une scène 3D, à partir de simples données pixelliques. L'utilisation d'attributs de chrominance, en plus de la luminance, permet bien souvent d'améliorer cette compréhension de la scène. Bon nombre d'applications requièrent une invariance des traitements vis-à-vis de l'illumination, que ce soit au niveau de ses variations ou de l'apparition d'ombres et de spécularités. D'autres algorithmes d'analyse d'images nécessitent au contraire de fortes variations de radiance à la surface des objets. Dans tous les cas, l'utilisation de modèles de réflexion et/ou d'information colorimétrique s'avère souvent très pertinente pour répondre à ces problématiques, que ce soit au niveau de la reconstruction 3D [Healey et Binford, 1988, Zhang *et al.*, 1999, Shashua, 1997, Basri et Jacobs, 2001, Aaron et Steven, 2005], de la séparation des composantes de réflexion [Tan et Ikeuchi, 2005], de la segmentation [Healey, 1989, Kravtchenko et Little, 1999, Gevers, 2003b, Maxwell et Shafer, 1994, Maxwell et Shafer, 1996, Maxwell, 1996] ou bien encore de la constance couleur [Finlayson, 1995, Finlayson, 1996, Finlayson et Schaefer, 2001, Lehman et Palm, 2001]. La complexité des scènes rencontrées nécessite toutefois l'utilisation de modèles de réflexion simples, requérant peu de paramètres, tels que le modèle lambertien, le modèle dichromatique, le modèle de Phong, ou bien encore le modèle de Torrance-Sparrow.

Dans la suite du manuscrit, nous nous intéressons plus particulièrement à deux problématiques. Dans un premier temps, il s'agit de traiter le problème de l'extraction de régions d'intérêt, c'est-à-dire des groupements de pixels connexes répondant à un même critère de similarité. Nous prouverons notamment tout l'intérêt de la couleur dans cette tâche et étudierons un exemple de segmentation de marquages transparents où l'information couleur permet de définir des attributs invariants vis-à-vis des phénomènes de transparence. D'autre part l'intérêt de ces attributs sera souligné par comparaison de notre approche avec une méthode de segmentation développée dans un espace couleur classique, parmi ceux décrits dans le chapitre 2.

Dans un second temps, nous étudions le problème de la robustesse du suivi de primitives vis-à-vis des variations d'illumination, en nous basant d'une part sur une analyse des modèles de réflexion décrits dans le chapitre 1, d'autre part sur les attributs d'invariance couleur.

Deuxième partie

**Apport de la couleur
pour l'extraction
de primitives**

INTRODUCTION

En vision par ordinateur, une primitive géométrique correspond à un élément caractéristique de l'image, tel que des contours, des droites, des coins, des régions, à savoir un regroupement de pixels connexes et proches au sens d'un critère. Elle constitue une information clé sur l'image, tout en réduisant l'importante masse de données que représentent les valeurs d'intensité des pixels. De ce fait, elles sont à la base de traitements plus évolués, tels que la reconnaissance d'objet ou bien encore l'asservissement visuel.

Cette partie du manuscrit s'intéresse tout particulièrement à l'extraction de primitives couleur (de type contours ou régions) par segmentation. Cette tâche consiste à créer une partition de l'image en régions. Dès lors que la géométrie de la ou des régions à détecter n'est pas connue *a priori*, toute méthode de reconnaissance de forme est illusoire. Alors que de nombreuses méthodes de segmentation appliquées à des images de luminance souffrent bien souvent de la présence de changements d'intensité dus aux ombres et aux courbures des surfaces, la segmentation couleur permet de discerner deux entités de même intensité, mais de teintes différentes. C'est pourquoi la littérature recense un très grand nombre de méthodes de segmentation dans les images couleur. Nous en dressons un état de l'art dans le chapitre 3. Les différentes techniques sont généralement classées selon trois catégories : les méthodes basées sur l'extraction de discontinuités (contours) dans l'image, les approches utilisant une classification des couleurs (à partir de l'espace couleur ou de l'histogramme) et enfin les méthodes de segmentation ascendantes ou structurales, qui partitionnent l'espace géométrique de l'image en régions homogènes. Toutefois, les travaux récents se penchent plutôt sur des approches mixtes, alliant contours et régions ou bien encore classification et régions. Cet état de l'art nous permettra de sélectionner le type d'approches s'adaptant au mieux à la problématique de détection d'une primitive dans l'image (et pas forcément une segmentation totale de l'image). Lorsque la scène comprend des objets aux caractéristiques colorimétriques bien distinctes, le processus de segmentation est assez direct. Cependant, la plupart des images naturelles sont constituées de nuances, de dégradés de couleurs qui, à l'évidence, compliquent cette tâche.

C'est par exemple le cas de certains marquages à l'encre (ou à la peinture), souvent translucides et dont la couleur dépend à la fois de la quantité d'encre appliquée et de la couleur du support sur lequel elle est imprimée. Pour peu que la quantité d'encre varie au sein d'un même marquage, la tâche de segmentation devient vite ardue. La segmentation peut être facilitée par un choix adapté des attributs couleur, comme cela a été évoqué dans le chapitre 2. Le chapitre

4 s'intéresse précisément à cette problématique, en proposant des attributs couleurs adaptés à l'analyse de ces phénomènes de transparence. Nous montrons leur pertinence dans une tâche de segmentation d'un marquage à l'encre, où l'algorithme utilisé se base sur une approche mixte classification/région. Des variantes de cet algorithme sont également proposées, afin de prendre en compte le cas de figure où les pixels associés au marquage à détecter ne sont pas forcément tous connexes (ou voisins).

Cette technique a été appliquée dans le cadre d'une étude sur le contrôle de la traçabilité de produits porcins, où il s'agit justement de détecter de manière automatique un marquage à l'encre sur des pièces de jambons. La quantité d'encre utilisée pour effectuer le marquage ainsi que la couleur de la couenne varient d'une image à l'autre, ce qui justifie l'utilisation d'attributs adaptés. Cette application fait l'objet du chapitre 5.

EXTRACTION DE PRIMITIVES COULEUR : ÉTAT DE L'ART EN SEGMENTATION D'IMAGES COULEUR

Ce chapitre dresse un état de l'art des méthodes de segmentation de primitives couleur de type région, que ce soit par la recherche de discontinuités entre les régions (les contours) ou d'homogénéités couleur. Attestant de la diversité des méthodes existantes, il nous permettra également de privilégier les techniques susceptibles de répondre à notre problématique de détection d'une région de forme inconnue qui fera l'objet des chapitres 4 et 5.



La segmentation, qui consiste en une partition de l'image en régions connexes, homogènes et disjointes est une étape réalisée en amont de traitements de plus haut niveau tels que la détection d'objets ou la reconnaissance de forme. Il s'agit soit d'extraire des contours dans l'image, soit de détecter une ou plusieurs régions, en se basant sur des critères et des attributs largement dépendants de l'application visée et du type d'images traitées. Basons-nous sur la définition initiale de la segmentation, donnée par Zucker [Zucker, 1976]. Soit f une image composée de T pixels p , \mathcal{P} un prédicat défini sur tout sous-ensemble de pixels connexes de f . Segmenter f consiste à réaliser une partition de f en N_R sous-ensembles de pixels $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{N_R}$, appelés régions, tel que :

- $\bigcup_{n=1}^{N_R} \mathcal{R}_n = f$.
- $\forall n = 1, \dots, N_R, \forall$ le couple de pixels $(p_1, p_2) \in \mathcal{R}_n \times \mathcal{R}_n, \exists$ un chemin de p_1 à p_2 dans \mathcal{R}_n .
- $\forall n = 1, \dots, N_R, \mathcal{P}(\mathcal{R}_n) = \text{Vrai}$.
- $\forall n \neq m, \mathcal{R}_n$ voisin de $\mathcal{R}_m \Rightarrow \mathcal{P}(\mathcal{R}_n \cup \mathcal{R}_m) = \text{faux}$

En d'autres termes, le résultat de la segmentation n'est autre qu'un ensemble disjoint de N_R régions où toute région \mathcal{R}_n pour $n = 1, \dots, N_R$ est constituée de pixels spatialement connexes. Tous les pixels constitutifs d'une région respectent un *prédicat d'uniformité*, prédicat basé dans le cas de la segmentation couleur sur l'homogénéité colorimétrique. Enfin, deux régions adjacentes doivent être suffisamment distinctes ce qui signifie implicitement que deux régions adjacentes trop ressemblantes doivent être fusionnées de manière à ne former qu'une seule et même région. À l'issue de la segmentation, une couleur aléatoire (une étiquette ou *label*) peut être affectée aux régions formées, de manière à visualiser la partition. Tandis que la quantification a pour but de réduire le nombre de couleurs de l'image tout en retranscrivant l'impression visuelle de l'image initiale (une même étiquette peut alors être affectée à plusieurs régions de l'image), la segmentation ne cherche qu'à partitionner l'image de manière très pragmatique dans le but de faciliter une analyse. Par contre, la quantification peut être utilisée comme filtrage préalable à une segmentation, comme dans [Deng *et al.*, 1999, Lucchese et Mitra, 2001, Phung *et al.*, 2005] par exemple.

La littérature recense un nombre incommensurable de méthodes de segmentation couleur. Elles peuvent être classées en trois catégories principales, selon qu'elles se basent sur la recherche de discontinuités locales entre les régions homogènes (les contours), sur l'extraction dans l'espace couleur de classes de pixels colorimétriquement homogènes (notons que des traitements supplémentaires sont alors nécessaires à l'obtention d'une segmentation telle que le décrit la définition de Zucker), ou encore sur la partition géométrique de l'image par analyse des propriétés spatiales et colorimétriques des pixels. Nous consacrons une section de ce chapitre à chacune de ces familles de techniques. En outre, une quatrième section regroupe les méthodes mixtes, dénommées ainsi dans le sens où la segmentation de l'image géométrique est améliorée par une association à une classification dans un espace d'attributs ou à une recherche de contours.

Dans cet état de l'art, nous insistons sur le fait que certaines techniques sont plus appropriées que d'autres à notre cadre d'étude où il s'agit non pas d'effectuer une segmentation globale de l'image mais de détecter une primitive couleur de forme inconnue et variable d'une image à l'autre. Soulignons que nous ne supposons aucune hypothèse quant à la structure de la primitive à détecter. Le lecteur intéressé par une explication plus détaillée et plus exhaustive des techniques de segmentation couleur pourra se référer par exemple aux revues bibliographiques plus ou moins détaillées de [Pal et Pal, 1993, Skarbek et A.Koschan, 1994, Cheng *et al.*, 2001, Navon *et al.*, 2005], aux ouvrages [Cocquerez et Philipp, 1995, Trémeau *et al.*, 2004] ou encore à la thèse [Raffy, 1999].

3.1 L'approche contour

Les techniques d'extraction de contours dans une image de luminance se basent sur la recherche des discontinuités, par le biais de méthodes dérivatives, morphologiques ou surfaciques. Les discontinuités d'une image couleur s'obtiennent de deux manières :

- soit de manière *marginale*, par l'analyse des discontinuités sur chacune des composantes couleur (C_1, C_2, C_3) de façon indépendante, de la même manière qu'une image de lu-

minance ; ces résultats marginaux sont ensuite combinés avec des opérateurs logiques de manière à obtenir un résultat monocomposante.

- soit de manière *vectorielle* en exploitant des grandeurs calculées de manière conjointe à partir des trois composantes colorimétriques.

3.1.1 Approches marginales

Dans une image couleur, contrairement au cas des images de luminance, la transition entre deux objets adjacents peut être caractérisée non seulement par une discontinuité dans le niveau de luminance, mais également par des changements brutaux de chrominance. Les dérivées premières (gradient) et/ou secondes (laplacien) sont exploitées dans ce but.

Le gradient. La dérivée du plan image C_i , $i = 1...3$, ou gradient ∇C_i se caractérise par ses deux coordonnées $G_x^{C_i}$ et $G_y^{C_i}$, mais également par sa norme ainsi que par son orientation. Les contours sont détectés dans chaque plan image par recherche des maxima du gradient. La figure 3.1 illustre l'obtention d'un gradient (figure 3.1b) à partir d'un contour schématisé sur la figure 3.1a. Une approximation du gradient est obtenue par convolution de l'image par des masques de type Sobel ou Prewitt [Prewitt, 1970], dans deux directions orthogonales. Néanmoins, il est préférable d'utiliser les masques de Deriche ou Shen-Castan [Shen et Castan, 1992], qui produisent généralement des résultats de meilleure précision. Le gradient marginal peut également s'obtenir en utilisant les outils de morphologie mathématique [Cocquerez et Philipp, 1995].

Les dérivées secondes. Le calcul des dérivées secondes, ou laplacien $\nabla^2 C_i(x, y)$ permet de détecter les contours ou d'améliorer une détection effectuée par analyse du gradient. La détection du passage par zéro du laplacien correspond à la fois aux maxima et aux minima du gradient, comme le montre la figure 3.1c. Ceci induit généralement la détection de contours fantômes. De plus, le laplacien agit comme un filtre passe-haut et introduit du bruit dans l'image, ce qui requiert l'application d'un filtre passe-bas (un filtre gaussien par exemple) préalablement au calcul du laplacien.

Du fait de ces artefacts, il est nécessaire de procéder ensuite à une étape de seuillage des contours, par hystérésis¹ par exemple [Canny, 1986]. L'ajustement de la valeur du seuil est très critique puisqu'il dépend du contenu de l'image, un seuil trop faible faisant apparaître de faux contours et à l'inverse un seuil trop élevé aboutissant à des non détections de contours. Afin de remédier à ce problème, certains auteurs les déterminent automatiquement de manière à ce que, quelle que soit l'image, un certain pourcentage des pixels de l'image soit détecté comme des contours.

La majorité des applications ne requiert qu'une seule image de contours, suffisamment satisfaisante d'un point de vue de la précision et des bonnes détections. Cette contrainte impose de fusionner les différents gradients. À cet effet, les opérateurs logiques ET/OU, très simples dans

¹Ce principe se base sur deux seuils, un seuil haut S_h et un seuil bas S_b . Les maxima locaux de valeur supérieure à S_h sont sélectionnés, ainsi que les maxima locaux de valeur supérieure à S_b et qui appartiennent à une composante connexe comprenant au moins une valeur supérieure à S_h . permet d'obtenir des contours mieux fermés que ceux obtenus en utilisant un seuil unique.

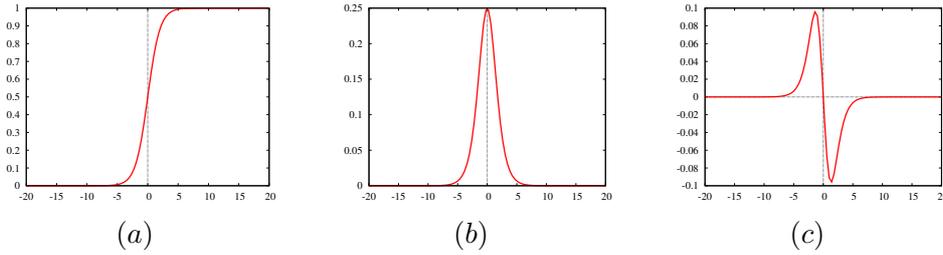


FIG. 3.1 – Illustration de la détection de contours. (a) Transition entre deux régions. (b) Dérivée première associée. (c) Dérivée seconde.

leur mise en œuvre, sont couramment utilisés. Ils ont l'avantage de pouvoir fusionner n'importe quel type de détecteur de contours. De plus, des détecteurs différents peuvent être utilisés sur chaque composante. La figure 3.2 illustre l'utilisation de ces opérateurs, sur l'image de la figure 3.2a. Un ET logique (figure 3.2b) permet d'éliminer les faux contours mais réduit la discrimination des contours significatifs ; un OU (figure 3.2c) améliore la détection, mais favorise les fausses détections. La fusion peut également s'opérer à l'aide d'opérateurs logiques dérivés de la logique floue. À chaque pixel p de l'image résultat est assignée une valeur qui dépend soit uniquement de la valeur des normes de chaque gradient marginal (valeur moyenne, pondérée ou non, ou maximum des 3 normes) soit à la fois des normes et des directions des gradients marginaux en p [Bloch, 1996]. Le lecteur pourra trouver une étude détaillée des approches de fusion dans [Bloch, 2003].

3.1.2 Approche multicomposante ou vectorielle

Du fait des fortes corrélations spatiales et chromatiques entre chacun des plans image, il est généralement plus satisfaisant de détecter les contours de manière globale. Nous présentons à ce propos le gradient vectoriel, valable dans tout espace couleur, et le gradient basé sur la teinte. Pour finir, nous abordons succinctement le cas des contours actifs couleur.

3.1.2.1 Gradient vectoriel

Le gradient multicomposante de [Zenno, 1986] et de [Lee et Cok, 1991] considère qu'une image exprimée dans l'espace (C_1, C_2, C_3) est représentée par un champ vectoriel \mathbf{C} . La matrice du gradient spatial \mathbf{D} de \mathbf{C} est alors définie par la matrice jacobienne de la fonction $(C_1(x, y), C_2(x, y), C_3(x, y))$ suivant les coordonnées spatiales x et y d'un point p de la manière suivante :

$$\mathbf{D}(p) = \begin{bmatrix} \frac{\partial C_1}{\partial x} & \frac{\partial C_1}{\partial y} \\ \frac{\partial C_2}{\partial x} & \frac{\partial C_2}{\partial y} \\ \frac{\partial C_3}{\partial x} & \frac{\partial C_3}{\partial y} \end{bmatrix} \quad (3.1)$$

Le calcul de la matrice symétrique semi-positive $\mathbf{D}^T \mathbf{D}(p)$ permet alors de caractériser le gradient. La racine carrée de la plus grande valeur propre de $\mathbf{D}^T \mathbf{D}$ correspond à l'amplitude du gradient. Le vecteur propre associé exprime la direction du gradient. Les contours s'obtiennent

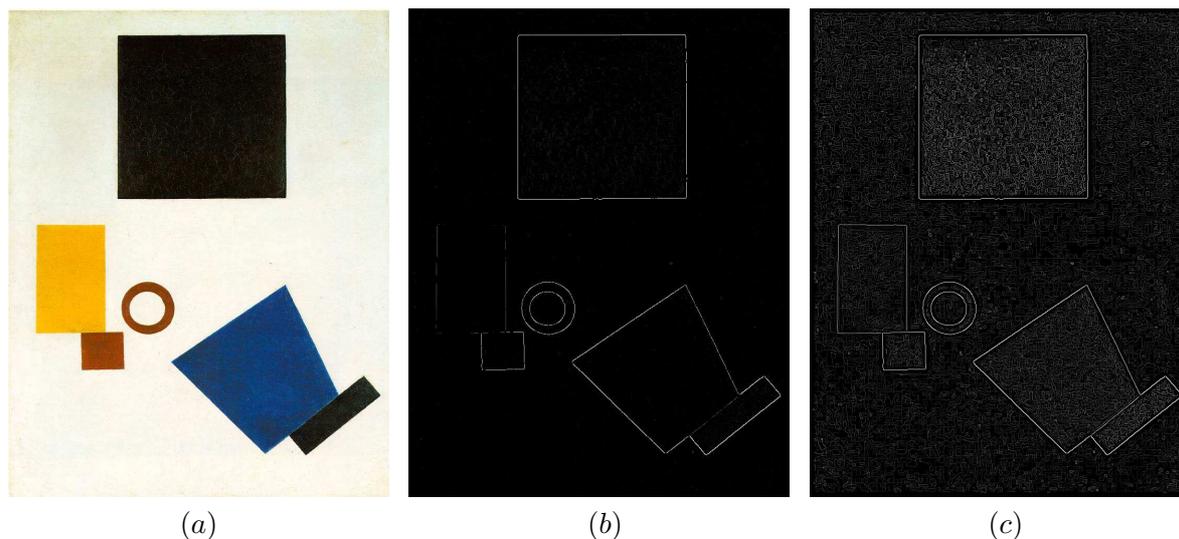


FIG. 3.2 – Illustration du gradient marginal (méthode de Deriche). (a) Image initiale (« Suprematism : self portrait in 2 dimensions », de Malevich). (b) Résultat d'un ET logique entre les trois résultats marginaux sur les composantes RGB. (c) Résultat d'un OU logique.

par la détection des maxima locaux de l'amplitude du gradient suivie d'un seuillage par hystérésis.

La section suivante présente un deuxième gradient multicomposante, basé sur le système perceptuel.

3.1.2.2 Le gradient basé sur le système perceptuel

Carron [Carron, 1995] propose deux méthodes de gradient spécifiques aux images couleur, qui se calculent dans un espace perceptuel de teinte T , saturation S et luminance L (voir le paragraphe 2.3.3 page 38). Leur originalité réside dans la prise en compte de la *pertinence de la teinte*. En effet, lorsque la couleur est fortement saturée, l'information de teinte est très fiable mais cette information perd de sa pertinence et devient très sensible au bruit dès lors que la couleur n'est pas suffisamment saturée. La pertinence de la teinte T peut être exprimée à l'aide d'un coefficient $\alpha(S)$, noté α qui tel que l'illustre la figure 3.3a, est une fonction croissante de la saturation, avec un point d'inflexion en S_0 . Le gradient de Carron, dans son expression générale, prend la forme suivante :

$$\mathbf{G} = \alpha \mathbf{G}^T + (1 - \alpha) \mathbf{G}^L + (1 - \alpha) \mathbf{G}^S \quad (3.2)$$

où \mathbf{G}^T , \mathbf{G}^L et \mathbf{G}^S sont les gradients calculés sur les plans couleur T , L et S respectivement. Ainsi, le gradient de teinte est privilégié lorsque celle-ci est pertinente, le cas inverse permettant de privilégier la saturation et la luminance. La figure 3.3b montre les résultats de la détection de contours avec le gradient basé sur la teinte et la saturation. En comparaison aux résultats obtenus avec les approches marginales (figures 3.2a et 3.2b), les contours s'avèrent plus marqués.

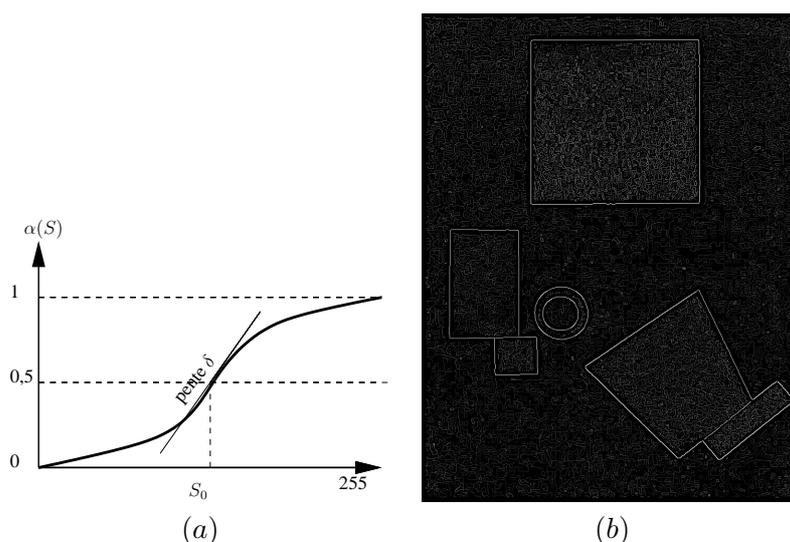


FIG. 3.3 – Gradient multicomposante basé sur la teinte. (a) Fonction de pertinence de la teinte, dont les paramètres sont la pente de l'inflexion δ et la position de cette inflexion S_0 . (b) Résultat d'une détection de contours sur l'image 3.2a.

3.1.2.3 Un cas particulier : les contours actifs

Les contours actifs ou déformables (en anglais *snakes*, ou *active contours*), dont le formalisme a été proposé par [Kass *et al.*, 1988] pour des images de luminance, sont particulièrement adaptés à la segmentation d'objet. Ils consistent à déformer un contour initial par minimisation d'énergie jusqu'à le faire converger vers le contour de l'objet à détecter. Grâce aux modèles déformables [Caselles *et al.*, 1993, Malladi *et al.*, 1994], mais surtout aux *level-sets* [Osher et Rudin, 1988], plusieurs objets peuvent être détectés sans connaissance *a priori* sur leur nombre. La littérature recense également un grand nombre d'approches spécifiques à la couleur [Sapiro, 1997, Zhu *et al.*, 1996, Chan *et al.*, 2000, Ballerini, 2003, Augereau *et al.*, 2005].

3.1.3 Discussion

La détection de contours repose sur la recherche de discontinuités dans l'image. Elle obtient généralement de bons résultats dans le cas d'images fortement contrastées, montrant des variations de luminance et de couleur bien nettes, et non graduelles entre les différents objets de l'image. Cependant, ces approches sont souvent très sensibles aux variations de texture et au bruit impulsionnel. De plus, l'obtention de lignes de contour continues est une réelle difficulté [Moghaddamzadeh et Bourbakis, 1997], et il s'avère souvent utile de recourir à des post-traitements tels que des méthodes de recherche du meilleur chemin entre deux extrémités d'un contour [Cocquerez et Philipp, 1995]. Excepté dans le cas d'images très simples (un objet de couleur très homogène sur un fond homogène), ou dans le cas d'objets de forme connue, l'utilisation exclusive de méthodes de détection de contours ne nous semble pas adaptée à une parfaite discrimination d'une zone d'intérêt vis-à-vis d'autres entités présentes dans l'image. Des traite-

ments d'analyse de forme peuvent être requis, or ceux-ci ne sont pas envisageables dans notre cadre d'étude où l'on ne suppose aucune hypothèse quant à la structure de la primitive.

Cette section 3.1 a montré que la segmentation d'images couleur peut être vue comme une recherche de discontinuités entre les objets. Une région est alors définie comme la zone intérieure d'un contour fermé. La section suivante expose quant à elle les méthodes basées sur la recherche d'homogénéités dans le système de représentation de la couleur : les méthodes de classification pixellaire.

3.2 La classification pixellaire

Le terme de *classification pixellaire* est utilisé ici pour regrouper les méthodes de partitionnement basées exclusivement sur les attributs couleur des pixels, sans exploiter d'informations spatiales issues de la représentation géométrique de l'image. Utiliser de la classification pixellaire pour effectuer une partition couleur revient à supposer :

- soit que les pixels d'une même région possèdent les mêmes caractéristiques colorimétriques, et forment des nuages de points couleur dans l'espace des couleurs 3D. C'est le cas de ce que l'on appelle les méthodes de *clustering* qui feront l'objet du premier paragraphe ;
- soit que la distribution des couleurs dans l'histogramme des couleurs forme des modes de forte densité de pixels correspondant aux classes présentes. Nous aborderons cette approche dans le deuxième paragraphe.

La classification ne peut répondre à tout type de problème de segmentation puisque tous les critères de définition de la segmentation explicités en introduction ne sont pas respectés dans l'absolu, notamment en ce qui concerne la loi de connexité. À titre d'exemple, la figure 3.4 montre globalement 5 classes dans l'espace des couleurs, tandis que l'image (celle de la figure 3.2) comprend sept régions distinctes. C'est pourquoi un algorithme basé sur une classification pixellaire est généralement suivi d'un étiquetage en composantes connexes. Certains auteurs effectuent la partition de l'espace d'attributs en s'aidant de la répartition géométrique des pixels. Ces méthodes de *classification spatio-colorimétriques* seront abordées dans le dernier paragraphe.

3.2.1 Le *clustering*

Les méthodes de *clustering*, empruntées au domaine de la statistique, consistent à partitionner l'ensemble des pixels de l'image en N_C classes dans l'espace des couleurs. Le nombre de classes N_C est généralement fixé par l'utilisateur, ce qui impose une connaissance *a priori* du contenu de l'image. La littérature recense de nombreuses méthodes de clustering, mais nous abordons ici trois ensembles d'approches majeures, à savoir les méthodes de type *centres mobiles*, l'algorithme d'ISODATA et pour finir les méthodes de classification floue.

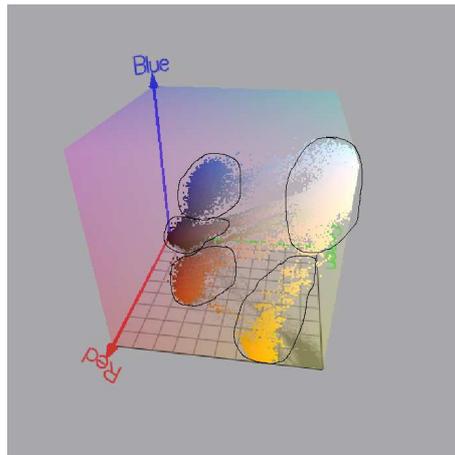


FIG. 3.4 – Représentation de l'image 3.2a dans l'espace de couleur RGB. On décompte 6 objets et le fond dans l'espace 2D, et 5 classes dans l'espace des couleurs.

Les algorithmes de type centres mobiles [Forgy, 1965] sont supervisés au niveau de leur initialisation. Pour chacune des classes, l'utilisateur sélectionne un pixel qui en devient le centre de gravité. Les pixels de l'image sont ensuite affectés à la classe dont le centre de gravité est le plus proche au sens de la distance colorimétrique considérée. Après affectation de l'ensemble des pixels de l'image, les centres de gravité sont remis à jour par calcul de la moyenne des pixels associés à la classe. S'il y a eu déplacement du centre de gravité, l'image est à nouveau balayée et ses pixels reclassés. L'algorithme prend fin lorsque les centres de gravité de toutes les classes sont suffisamment stabilisés. La littérature cite de nombreuses variantes de ces méthodes, parmi lesquelles :

- les *nuées dynamiques* ; cette méthode diffère des centres mobiles de manière peu sensible, puisqu'une classe n'est plus représentée par son centre de gravité mais par l'observation la plus proche de ce centre de gravité ;
- la méthode des *K-means* [Hartigan, 1975] s'avère la plus populaire des techniques de clustering. Dans ce cas de figure, chaque nouvelle affectation d'un pixel à une classe induit la mise à jour du centre de gravité. Elle est par exemple utilisée dans [Lucchese et Mitra, 2001], après diffusion anisotropique² de la chromaticité et de la luminosité ou dans une application d'analyse de documents dans [Leydier *et al.*, 2004].

Toutes ces méthodes visent à minimiser la *variance intraclasse*, basée sur la somme des distances au carré des pixels par rapport à leur centre de classe. Le caractère itératif de cette approche mais également la nécessité de fixer le nombre de classes représentent ses deux principales faiblesses. L'algorithme ISODATA répond à cette dernière limitation.

L'algorithme ISODATA (pour Iterative Self-Organizing Data Analysis Techniques A), proposé par [Takahashi *et al.*, 1995], permet de renoncer au caractère supervisé des méthodes précédentes en s'affranchissant de la connaissance du nombre de classes. Celui-ci évolue au

²La diffusion anisotropique est une technique de lissage qui est encouragée dans les régions uniformes et inhibée au niveau des contours.

cours des itérations selon un procédé de division de classes dans l'espace des couleurs. Dans un premier temps, l'ensemble des attributs de couleur des pixels présents dans l'image forme une seule et même classe (ou *cluster*). Le processus de division et fusion des classes évolue en se référant à deux seuils fixés par l'utilisateur. Le premier tolère une certaine hétérogénéité des pixels d'une classe, au-delà de laquelle la classe est divisée. Le second seuil permet de fusionner les classes dès lors que les centres de gravité de deux classes deviennent trop proches. De plus, les classes contenant très peu de pixels sont éliminées. La détermination de tous les seuils et paramètres énoncés dans l'algorithme reste bien sûr un problème, puisqu'elle est généralement empirique et dépend du type d'application ainsi que du type d'image utilisé. De plus, son caractère itératif est pénalisant d'un point de vue du temps de calcul. L'algorithme d'ISODATA n'en reste pas moins intéressant par sa capacité d'auto-organisation et par sa flexibilité.

La classification floue permet de considérer une classe de pixels comme un ensemble flou, où la vraisemblance de l'appartenance d'un pixel p à une classe k est représentée par une *fonction d'appartenance* $\mathcal{A}_k(p)$, c'est-à-dire par sa probabilité d'appartenir à la classe k et à aucune des autres classes. Bezdeck [Bezdeck *et al.*, 1984] proposa l'algorithme de *fuzzy c-means*, technique itérative non-supervisée de classification, qui a été adaptée par la suite au cas de la couleur [Lim et Lee, 1990]. La fonction d'appartenance du point p est calculée à partir des distances de ce point vis-à-vis de toutes les autres classes. Dans un premier temps, le nombre de classes et leurs centres de gravité sont fixés par l'utilisateur. Ensuite, les valeurs d'appartenance $\mathcal{A}_k(p)$ sont mises à jour de manière récursive jusqu'à leur stabilisation. Le partitionnement est alors déterminé par minimisation d'une fonction de coût, qui dépend à la fois de la distance colorimétrique par rapport au centre de la classe et de la fonction d'appartenance du pixel à cette même classe. Très souvent l'algorithme de fuzzy C-means est exploité dans un but d'affinement d'une segmentation. Dans [Lim et Lee, 1990], l'emploi de ce type d'approches permet de lever les ambiguïtés de classification de certains pixels isolés, après une analyse d'histogrammes. Ce type d'approches a également été exploité dans [Yang *et al.*, 2002] et [Chen et Lu, 2002]. Dans ce dernier article, une notion de taille de classe est intégrée dans l'expression de la fonction d'appartenance. Comme dans tout algorithme de classification, l'étape d'initialisation est critique. C'est pourquoi [Kim *et al.*, 2004] propose d'améliorer l'initialisation de l'algorithme de fuzzy c-means en se basant sur l'hypothèse selon laquelle les couleurs les plus dominantes et vives n'appartiennent pas à un même *cluster*. Ainsi, l'utilisation des concepts de logique floue permet d'éviter de classer les pixels dans une région de manière binaire de type *appartient* ou *n'appartient pas*. Elle pondère une appartenance à une région, et préserve ce flou le plus longtemps possible avant qu'une décision ne soit prise. En contrepartie, l'ajustement des paramètres de la fonction d'appartenance est une tâche délicate.

Les approches de classification statistiques expriment le problème de classification en termes probabilistes, où la classe est supposée suivre une distribution spécifique dans l'espace couleur, classiquement une gaussienne, et où chaque pixel correspond à une variable aléatoire. On trouvera des explications plus approfondies dans [Raffy, 1999]. L'approche bayésienne, la plus répandue, a récemment été exploitée dans le cadre d'applications en imagerie microscopique [Lezoray et Cardot, 2002a] et pour effectuer la segmentation de la couleur de peau [Phung *et al.*, 2005].

3.2.2 La classification par analyse d'histogrammes

Les techniques basées sur l'analyse d'histogrammes reposent sur l'hypothèse selon laquelle les objets se manifestent comme des groupements dans l'espace de mesure, les régions de même couleur laissant dessiner des modes dans les histogrammes des composantes couleur. L'histogramme peut être utilisé dans sa version tridimensionnelle, puisque cette structure de données représente au mieux la distribution des pixels d'une image couleur. Il est toutefois plus usité dans ses versions bidimensionnelle et unidimensionnelle pour des raisons évidentes de complexité calculatoire. Ces différentes approches sont détaillées ci-dessous.

Analyse d'histogrammes multidimensionnels. Étant donné l'aspect tridimensionnel de l'information colorimétrique, il paraît naturel de procéder à des analyses d'histogrammes 3D. Ce type de méthodes aboutit à la détection des groupes de pixels, et à des maxima de densité correspondant à des classes dans l'espace couleur considéré. Là encore, comme pour les méthodes de classification précédemment étudiées, seuls les critères colorimétriques interviennent. À cause du fort coût en temps de calcul qu'elles représentent, les méthodes d'analyse des histogrammes 3D sont assez rares dans la littérature. Généralement, l'extraction des modes s'effectue par binarisation-étiquetage (voir par exemple [Postaire, 1987, Raffy, 1999]), par extraction récursive [Raffy, 1993, Raffy, 1999] ou à l'aide de techniques issues de la morphologie mathématique [Park *et al.*, 1998].

Analyse d'histogrammes bidimensionnels. Certains auteurs n'hésitent pas à ignorer l'un des trois plans chromatiques de manière à réduire notablement la quantité de données à traiter. La figure 3.5 représente un exemple d'histogramme 2D, celui du plan couleur GB de l'image 3.2a. Un des avantages de l'histogramme 2D est de pouvoir être assimilé à une image en luminance, sur laquelle des algorithmes de traitement d'images classiques sont utilisables. Les travaux précurseurs sont dus à Holla [Holla, 1982]. L'auteur transforme l'espace RGB en espace luminance-chrominance, les composantes de chrominance étant définies par les paires de couleurs opposées Rouge-Vert (RG) et Bleu-Jaune (YB). Les classes y sont alors définies par des aires. Dans [Kurugollu *et al.*, 2001], les auteurs développent une méthode de multiseuillage d'histogrammes 2D basée sur la théorie de [Koonty *et al.*, 1976]. À partir des N composantes de l'image sont obtenus C_N^2 histogrammes 2D. Après recherche des pics significatifs, les pixels sont affectés à une classe suivant une *force de gravité* d'une case de l'histogramme par rapport à chaque pic. Dans une optique similaire, Lezoray *et al.* [Lezoray et Cardot, 2002b] filtrent les histogrammes bidimensionnels par morphologie mathématique afin d'obtenir un centre associé à chacune des classes présentes. Ils mettent ensuite en œuvre un partitionnement de l'histogramme en régions de Voronoï³.

Analyse d'histogrammes monodimensionnels. Les méthodes d'analyse d'histogrammes 1D ont intéressé et intéressent encore de nombreux auteurs. En guise d'exemple, les figures 3.6a, 3.6b et 3.6c montrent les histogrammes 1D de l'image 3.2a, où différents modes M_{i_j} apparaissent dans les histogrammes RGB . La figure 3.7 représente les différentes classes associées

³les régions de Voronoï sont décrites dans la section 3.3.2

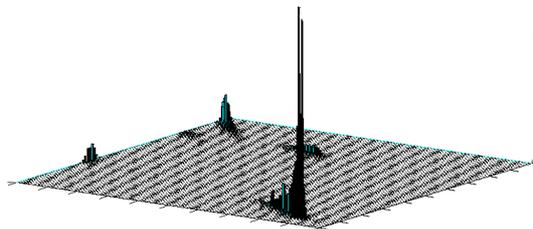


FIG. 3.5 – Histogramme bidimensionnel du plan GB de l'image 3.2a.

à ces modes, après la fixation de seuils. Certes, travailler sur l'histogramme 1D permet de réduire la complexité d'un traitement dans l'espace couleur 3D, mais la projection de l'espace 3D sur un sous-espace de dimension inférieure contribue à la perte d'information, notamment la perte de modes.

Les techniques d'analyse d'histogrammes 1D se différencient par les techniques d'extraction des modes, le critère d'arrêt, mais également le nombre de composantes couleur utilisées. Nous abordons ici ces trois points.

- *Extraction des modes.* La phase d'extraction des modes débute généralement par la recherche du mode le plus significatif dans les histogrammes considérés. Une fois les pixels associés à ce mode écartés du processus, l'algorithme détecte le second pic le plus significatif, et ainsi de suite. Plusieurs critères de sélection de modes ont été proposés dans la littérature. Certains d'entre eux prennent en compte l'amplitude et la largeur des modes [Ohlander *et al.*, 1978, Otha *et al.*, 1980], la population du mode et la population totale de l'histogramme [Tominaga, 1986] ou bien encore l'entropie [Guo *et al.*, 1998]. Plus récemment, [Delon *et al.*, 2005] emploie le principe de Helmholtz. L'histogramme est découpé itérativement suivant l'analyse de critères d'unimodalité, d'uniformité et de monotonie. Les méthodes multi-échelle, comme le *scale space filtering* [Schettini, 1993], s'avèrent également pertinentes. Ce type de technique a en effet démontré toute son efficacité pour ce qui est de la décomposition d'histogrammes en une somme de distributions gaussiennes. Le traitement consiste alors à convoluer l'histogramme avec des filtres gaussiens d'écart-type variables, de manière à obtenir la partition d'histogramme la plus adaptée.

Lorsque l'ensemble des modes significatifs est détecté, il est admis que certains pixels restants (associés à du bruit par exemple) ne sont constitutifs d'aucune classe dans l'image. Il est alors nécessaire de fixer un critère d'arrêt de l'algorithme itératif.

- *Critères d'arrêt.* Là encore, le processus itératif de recherche de modes peut être stoppé lorsqu'il n'y a plus de pic suffisamment représentatif c'est-à-dire lorsque toutes les régions sont supposées homogènes [Ohlander *et al.*, 1978]. Le critère d'arrêt peut également intégrer une notion de taille de région [Tominaga, 1986]. Dans cet article, la classification s'arrête soit lorsque l'aire d'une région est suffisamment petite par rapport à la taille initiale de l'image, soit

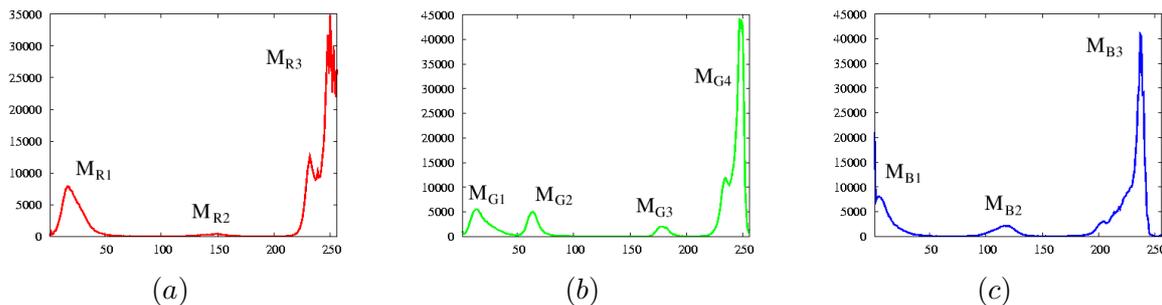


FIG. 3.6 – Histogrammes unidimensionnels associés à l'image de la figure 3.2a. (a) Composante R. (b) Composante G. (c) Composante B.

lorsque plus aucun histogramme ne possède de pics significatifs. De plus, les pixels non classés sont fusionnés avec une région voisine de même couleur.

- *Nombre de composantes.* Suivant l'application, l'étape de classification ne requiert pas nécessairement l'utilisation de plusieurs composantes couleur. Généralement, lorsque l'on a affaire à un type d'images bien déterminé, une seule composante choisie méticuleusement, peut suffire au traitement. Par exemple, Celenk [Celenk, 1990] utilise le *discriminant linéaire de Fischer* pour trouver une ligne dans l'espace tridimensionnel LHC permettant une bonne séparation des modes. Un seul histogramme est alors nécessaire. D'un point de vue plus applicatif, Lin et Chen [Lin et Chen, 1991] s'intéressent au suivi de route par segmentation d'images où il s'agit de diviser une image de paysage en deux classes : route et non route. Ils utilisent un seul attribut V_0 dépendant de l'intensité et de la saturation. Un simple seuillage de l'histogramme de V_0 , par détection des modes, suffit alors à séparer efficacement les deux classes, et ce malgré la présence éventuelle de parties ombragées sur la route.

Nous avons vu pour l'instant que les méthodes de classification exploitent uniquement l'information couleur mais aucune information quant à la localisation spatiale des pixels d'une même classe, ce qui amène souvent à opérer en aval un étiquetage en composantes connexes de façon à segmenter l'image. Certaines techniques permettent cependant de prendre en compte l'aspect spatio-colorimétrique dans la phase de classification.

3.2.3 Classification spatio-colorimétrique

La classification spatio-colorimétrique peut être vue comme un enrichissement de la classification couleur à l'aide d'attributs extraits de la géométrie de l'image. Ferri *et al.* [Ferri et Vidal, 1992] incluent des informations de forme et de surface dans la classification. Utilisant une méthode basée sur la technique des k plus proches voisins, un vecteur prenant en compte les composantes du pixel à traiter et de ses quatre voisins est créé. Huang *et al.* [Huang, 1992] résolvent le problème de la segmentation à l'aide de réseaux de neurones de Hopfield par minimisation d'une fonction de coût. Celle-ci vise à favoriser l'assignement des pixels adjacents à une même classe, tout en évitant l'attribution d'un pixel à plusieurs classes. Campadelli *et al.* [Campadelli *et al.*, 1997] ont adapté cette méthode aux images couleur. Il s'agit alors d'ini-

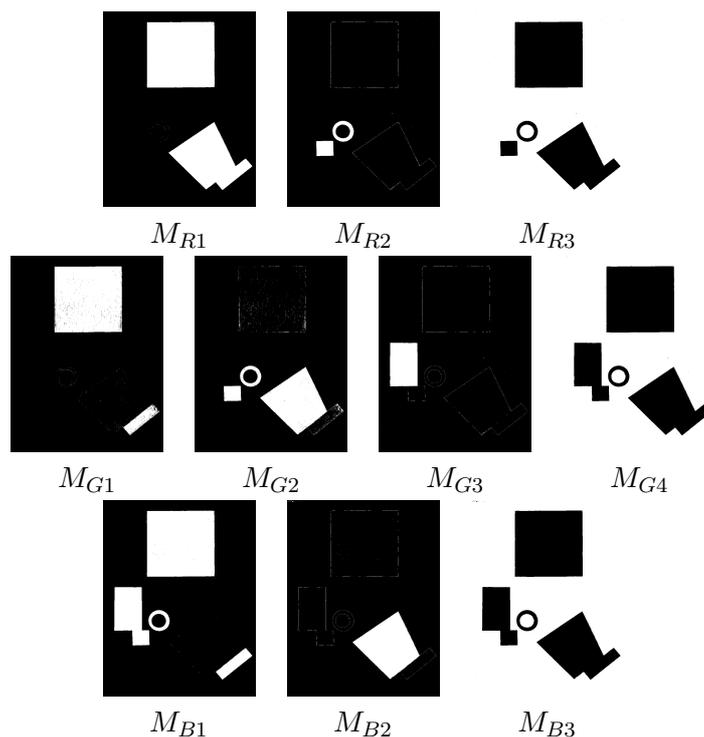


FIG. 3.7 – Exemples de classes extraites par seuillage d’histogrammes 1D.

tialiser un réseau de Hopfield en considérant le nombre de classes obtenues par analyse de l’histogramme. Notons également que des informations géométriques peuvent être introduites dans des techniques de type classification floue [Noordam et Broek, 2000].

Ensuite, la classification spatio-colorimétrique peut passer par la création de structures de données adaptées. La *pyramide des degrés de connexité couleur* ou PDCC [Fontaine, 2001] permet une analyse multi-échelle de l’image, en calculant un degré de connexité (basé sur le calcul des probabilités de cooccurrence) pour tout intervalle de couleur possible dans une image bichromatique. Cette structure de données est représentée sur la figure 3.8. Chaque cellule est repérée par ses coordonnées suivant les deux composantes bichromatiques C_1 et C_2 ainsi que par la taille l de l’intervalle de couleur considéré. Ainsi les plans proches de la base de la pyramide regroupent des cellules contenant les degrés de connexité des ensembles de pixels dont la couleur appartient à de petits intervalles de couleur. Le degré de connexité couleur y est généralement faible car la probabilité de cooccurrence d’un couple de couleurs qui appartient à un petit intervalle de couleur est faible.

En ce basant sur un principe similaire de structure de données spatio-colorimétrique multi-échelle, Cheng *et al.* [Cheng *et al.*, 2002, Cheng et Li, 2003] proposent le concept d’*homogramme*. Au lieu de calculer un degré de connexité, cette structure prend en compte un vecteur d’homogénéité floue d’un pixel par rapport à ses huit voisins. Les classes sont extraites par simple seuillage de cette structure de données. Notons cependant que ces approches de classification, même si elles intègrent des informations de connexité spatiale, ne satisfont pas forcément la définition de la segmentation donnée en introduction de ce chapitre.

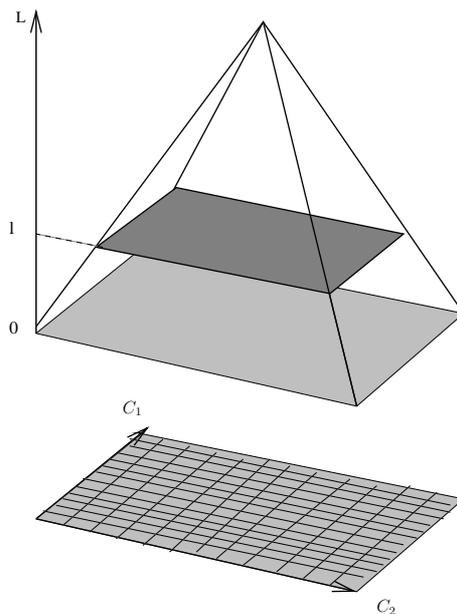


FIG. 3.8 – Pyramide des connexités.

3.2.4 Discussion

Les méthodes de classification pixellaire se basent soit sur la répartition des couleurs de pixels dans un espace couleur, pour en extraire des classes par des approches de clustering, soit sur la distribution des couleurs dans les histogrammes. Les méthodes de clustering offrent la simplicité du point de vue de l'implantation et un coût en temps de calcul très intéressant. Par contre, elles nécessitent généralement une supervision de l'utilisateur, que ce soit au niveau du choix du nombre de classes (c'est le cas des algorithmes de k-means), qui bien évidemment change d'une image à l'autre, ou dans la détermination de seuils et de divers paramètres de contrôle (c'est le cas de l'algorithme d'ISODATA). La classification par analyse d'histogrammes permet en général de s'affranchir de la connaissance *a priori* quant au contenu de l'image, et donc du nombre de classes. Néanmoins, les méthodes de classification par analyse d'histogrammes multidimensionnels sont très peu exploitées du fait des temps de calcul prohibitifs. Quant à l'histogramme 1D, sa facilité d'utilisation en fait un outil très populaire, malgré une séparation des modes souvent peu satisfaisante sur des images réelles, entraînant généralement des sous-segmentations. En effet, une classe séparable dans un espace couleur 3D ne l'est plus forcément après projection suivant une des composantes.

Dans le cas d'applications où une seule primitive couleur doit être détectée, la détection par classification est d'autant plus efficace que la primitive région est très homogène d'un point de vue couleur. Cela suppose que toutes les couleurs détectées soient constitutives de la primitive à détecter, et donc particulièrement dissemblables des caractéristiques de toute autre classe présente dans l'image. Dans ce contexte précis, la classification peut permettre l'extraction d'une primitive de forme inconnue. Dans le cas contraire, des traitements supplémentaires sont requis, relevant de l'analyse de forme ou de texture.

Les méthodes de classification ne prennent en compte que l'aspect colorimétrique des pixels et aucunement leur répartition spatiale, ce qui constitue leur principal point faible. D'autre part, elles aboutissent généralement à une sur- ou sous-segmentation de l'image. Par conséquent, elles sont très souvent utilisées dans une étape de pré-segmentation. Les méthodes de segmentation effectuées dans l'espace géométrique de l'image, et non uniquement dans l'espace couleur, peuvent résoudre ces différents problèmes.

3.3 Segmentation dans l'espace géométrique de l'image

En plus de considérer l'information colorimétrique des pixels, les méthodes de segmentation présentées dans cette section prennent en compte de manière très forte les interactions spatiales entre les pixels. La définition de la segmentation donnée page 49 se vérifie alors de manière exacte. Les différentes approches peuvent être classées en deux catégories :

- les *méthodes ascendantes*, consistant à faire croître une région à partir d'un pixel germe, à l'aide d'un critère d'homogénéité appliqué à son voisinage ;
- les *méthodes structurelles*, telles que la fusion, la division-fusion de régions, qui font appel à des structures de représentation des relations entre les régions, telles que les graphes.

3.3.1 Approche ascendante ou croissance de région

La méthode de *croissance de région* génère des régions uniformes en agglomérant les pixels les uns aux autres. L'algorithme commence par la sélection de N pixels *germes*, représentant N régions, de manière supervisée ou non. Lorsqu'un pixel voisin d'un germe satisfait un *critère de similarité*, il est aggloméré et la région se voit ainsi agrandie. Le voisinage du pixel nouvellement aggloméré est à son tour analysé et ainsi de suite.

Généralement, les résultats de la segmentation sont fortement dépendants de la sélection des germes, de la même manière qu'ils sont dépendants de l'ordre de traitement des régions, et du critère d'homogénéité sélectionné pour effectuer la croissance. C'est sur ces deux points que se différencient les méthodes de segmentation.

◦ *Le choix des germes*. Il peut s'effectuer par balayage vidéo de l'image ou de manière aléatoire. Toutefois, une approche plus intéressante consiste à rechercher des minima de l'image des contours. Les régions de valeur minimum correspondent aux valeurs d'homogénéité maximale. Les germes y sont alors sélectionnés. En se basant sur cette idée, [Ikonomakis *et al.*, 2000] détecte les germes dans des régions de faible variance de manière itérative. Plus récemment, Shih *et al.* [Shih et Cheng, 2005] proposent une sélection automatique de germes en se basant sur le fait qu'un germe doit avoir une forte similarité avec ses voisins (au sens de la 8-connexité).

◦ *Le critère d'homogénéité*. Une fois les germes sélectionnés, il reste à contrôler la croissance des régions en fixant les *critères d'homogénéité* adaptés. Le critère le plus simple se base sur une approche locale [Chassery et Garbay, 1984, Colantoni, 1998], par le biais de la distance colorimétrique entre le pixel à classer et son pixel voisin tout précédemment classé. Toutefois, ce critère se comporte mal sur des zones de dégradé, qui sont considérées comme homogènes.

Les résultats sont améliorés en calculant la distance colorimétrique entre le pixel à classer et l'ensemble du voisinage (les voisins ayant été classés dans la région courante). L'utilisation d'une approche globale peut également être judicieuse [Colantoni, 1998, Deng *et al.*, 1999], en comparant le pixel à traiter avec la moyenne de la région en cours de construction. Il existe de nombreux critères d'homogénéité couleur, basés sur le nombre de pixels d'une région, et sur des mesures de distance colorimétrique entre pixels constitutifs de cette région. Pour de plus amples détails, on pourra se référer à [Colantoni, 1998] par exemple. Moghaddamzadeh *et al.* [Moghaddamzadeh et Bourbakis, 1997] utilisent quant à eux un critère d'homogénéité flou. En effet, une région peut être définie comme une *collection de pixels connexes ayant à peu près la même couleur*, définition qui introduit des termes empruntés au langage flou. Les auteurs définissent une fonction d'appartenance basée sur la distance euclidienne entre deux couleurs. Le lecteur trouvera dans [Maeda *et al.*, 2000] un autre exemple de croissance de régions floues. Enfin, le *prédicat hybride* est le résultat d'une combinaison logique entre plusieurs des prédicats précédents, en fonction des attentes de l'utilisateur quant à la souplesse du prédicat. Par exemple, Trémeau *et al.* dans [Trémeau et Borel, 1997, Trémeau, 1998] préconisent l'utilisation de plusieurs critères basés sur le calcul de différentes distances colorimétriques : entre le pixel courant et le pixel précédemment traité, entre le pixel et le voisinage précédemment traité et enfin entre le pixel et la région en cours de formation.

Dans le cas d'une détection de primitive de type région, la principale difficulté liée à l'utilisation d'une méthode de croissance de région est de sélectionner des germes associés à la région à détecter, et exclusivement à celle-ci. Dans le cas général, la sélection des germes doit alors passer par un prétraitement autre que ceux présentés dans cette section. Une fois les germes correctement sélectionnés, les techniques de croissance de régions sont tout à fait adaptées à la détection d'une région de forme inconnue.

3.3.2 Approche structurelle : division-fusion de régions

Les techniques de segmentation par division-fusion débutent par une sur-segmentation de l'image en régions disjointes, puis accomplissent une fusion de ces régions de manière à ce qu'elles satisfassent un critère d'homogénéité colorimétrique. La plupart d'entre elles requiert l'utilisation de structures de données particulières, ce qui justifie leur appellation de méthodes *structurelles*. À ce titre, le tétra-arbre [Horowitz et Pavlidis, 1974] est une structure de données très commune de part sa simplicité et son faible temps de calcul. La technique de division peut être représentée sous la forme d'un *tétra-arbre* dans lequel la racine correspond à l'image entière et un nœud possède exactement quatre nœuds fils. La figure 3.9a illustre le principe de divisions successives d'une image en quadrants. Le lecteur pourra trouver un exemple de mise en œuvre dans [Gevers et Groen, 1991]. Les algorithmes basés sur l'analyse d'un tétra-arbre peuvent être résumés de la manière suivante :

1. L'image est divisée en quatre quadrants disjoints, représentant chacun une région
2. Toute région ne satisfaisant pas un critère d'homogénéité est divisée de la même manière en quatre et ainsi de suite, jusqu'à ce que chaque sous bloc respecte le prédicat d'homogénéité.
3. Deux régions voisines possédant des caractéristiques colorimétriques identiques sont fusionnées. Un graphe d'adjacence de régions permet alors de représenter les relations entre régions.

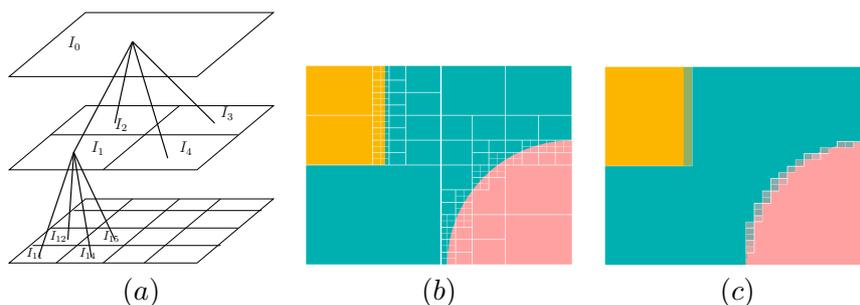


FIG. 3.9 – Algorithme de division-fusion. (a) Schéma de la division de l'image en quadrants. (b) Exemple d'une image avec sa division en quadrants. (c) Image après fusion des régions.

4. Le processus itératif s'arrête dès qu'il n'existe plus de couples de régions dont l'union respecte le prédicat d'uniformité.

On peut voir sur la figure 3.9b l'exemple d'une image après divisions en quadrants, et sur la figure 3.9c, la même image après fusion de régions similaires. De cet algorithme général découlent d'autres algorithmes spécifiques qui se différencient surtout par les critères de fusion et les prédicats d'uniformité.

Le principal défaut de l'utilisation du tétra-arbre est son incapacité à s'ajuster aux structures réelles des régions présentes dans l'image, de part la nature rectiligne et rigide des régions formées. L'utilisation de la triangulation de Delauney [Gevers et Kajcovski, 1994] permet d'adapter le partitionnement de l'image à l'agencement des régions et faire en sorte que les partitions suivent les lignes des frontières dans l'image. C'est également le cas du partitionnement par *diagramme de Voronoi** [Panjwani et Healey, 1995, Imelinska *et al.*, 2000].

Notons que la fusion des régions ne s'opère pas nécessairement après un algorithme de division de l'image, mais peut être accomplie après un algorithme de segmentation ayant provoqué une sur-segmentation. Les approches de fusion se basent généralement sur l'analyse du graphe d'adjacence de régions ou GAR. Les régions y sont représentées par les sommets du graphe et l'information d'adjacence entre régions est symbolisée par les arêtes. On trouvera différents types de fusions dans [Schettini, 1993]. Dans [Lezoray et Elmoataz, 2003], les auteurs simplifient le graphe d'adjacence de régions à l'aide d'un filtrage non-linéaire de manière à améliorer la fusion de régions. Enfin, l'automatisation non-supervisée de la tâche de fusion a été améliorée par [Navon *et al.*, 2005]. Dans un premier temps, l'image est fusionnée entièrement et toutes les étapes de ce processus ainsi que l'évolution des critères d'hétérogénéité associés aux régions en formation sont sauvegardées. Les trop fortes augmentations d'hétérogénéité permettent de déterminer à quelle étape de la fusion deux régions différentes ont été fusionnées à tort, et ainsi d'aboutir aux seuils locaux entre ces deux régions.

3.3.3 Discussion

Les méthodes de segmentation dans l'espace géométrique, ou méthodes basées région, prennent en compte à la fois l'aspect colorimétrique des pixels et leur répartition spatiale. Dans

ce sens, elles répondent parfaitement aux critères imposés par la tâche de segmentation (voir la définition de cette tâche page 49).

La principale difficulté réside dans le choix des critères de comparaison ou de division d'une région, qui doit être adapté à l'image considérée. Les algorithmes de croissance de régions requièrent la sélection de multiples seuils. Les résultats sont fortement dépendants de l'initialisation des germes et du critère d'homogénéité employé. En outre, du fait de sa nature séquentielle, ce type d'approches nécessite un temps de calcul important.

Les méthodes de division-fusion requièrent des structures de données souvent complexes et s'avèrent donc elles aussi coûteuses en temps de calcul. Toutefois, l'aspect le plus négatif de ces techniques reste celui de la géométrie de la partition engendrée, qui ne respecte pas parfaitement les contours de l'image, le pire exemple étant bien sûr celui de la division en quadrants. L'utilisation du diagramme de Voronoï ou de la triangulation de Delauney permet d'améliorer sensiblement les résultats mais au prix de calculs plus coûteux. Les approches de type fusion ou division-fusion ne nous semblent pas les mieux adaptées à l'application d'extraction de primitives précises dans l'image. En effet, cela ne justifie pas la segmentation globale de l'image. Cela nécessiterait d'utiliser en aval une technique de reconnaissance de formes.

Les différents types de segmentations abordés jusqu'alors, l'approche contours, la classification et la segmentation basée région, peuvent bien entendu être utilisées de manière conjointe. C'est le sujet de la prochaine section.

3.4 Méthodes mixtes

Parmi les approches dites *mixtes*, différencions d'une part les techniques se basant sur l'utilisation conjointe d'approches région et d'approches contour et d'autre part les approches alliant classification et segmentation géométrique.

3.4.1 L'association d'approches de type région et de type contour

Régions et contours ne sont pas dissociables. L'image segmentée, qui est en quelque sorte une carte des similarités entre régions, doit être conforme à l'image des contours, qui elle, discrimine les discontinuités entre ces régions.

Pour la plupart des techniques mixtes proposées, l'information de contour est introduite dans le but d'améliorer une approche de segmentation en régions [Zhu *et al.*, 1996, Saber *et al.*, 1997, Moghaddamzadeh et Bourbakis, 1997, Zugaj et Lattuati, 1998, Fan *et al.*, 2001]. Une coopération plus conjointe est obtenue à l'aide des approches de *ligne de partage des eaux* (ou *watershed* en anglais). Celles-ci se basent sur la description des images en termes géographiques. Une image peut en effet être perçue comme un relief si l'on associe l'intensité de chaque point à une altitude. La notion de crête est alors associée aux fortes intensités et la notion de vallée aux faibles intensités. On définit la ligne de partage des eaux comme étant la crête formant la limite entre deux vallées ou bassins versants. La dualité entre bassins versants et ligne de partage des eaux revient à effectuer une segmentation en régions à partir d'informations de contours (voir

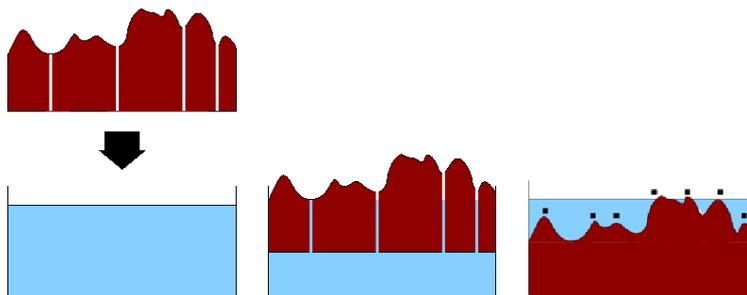


FIG. 3.10 – Illustration de la méthode des lignes de partage des eaux : technique de l’immersion.

paragraphe 3.1). L’efficacité de la méthode de ligne de partage des eaux dépend fortement des marqueurs de départ, c’est-à-dire des *minima* associés aux vallées. Sans traitement préalable, tel que l’élimination des *minima* locaux dus au bruit dans l’image, on obtient le plus souvent une sur-segmentation de l’image. Le premier algorithme d’inondation a été proposé par [Vincent et Soille, 1991]. L’algorithme de partage des eaux fut adapté plus tard au cas spécifique des images couleur [Meyer, 1992, Saarinen, 1994]. Le lecteur pourra également se référer à [Colantoni, 1998] où la méthode de l’immersion, illustrée par la figure 3.10, est appliquée sur des images de tissus, ou bien encore à [Cocquerez et Philipp, 1995] où cette même technique est utilisée pour effectuer une segmentation floue. Plus récemment, l’approche de montée des eaux a fait l’objet d’améliorations, au niveau de la détection des contours [Angulo et Serra, 2003, Kazanov, 2004] ou par l’intégration d’attributs liés à la physique [Nikolaev et Nikolayev, 2003].

3.4.2 L’association d’une classification dans l’espace des couleurs et d’une segmentation dans l’espace géométrique

Dans le cas d’images réelles, présentant par exemple des variations d’illumination, le nombre de classes issues d’une méthode de classification est généralement supérieur au nombre de régions réelles dans l’image. C’est pourquoi certains auteurs préconisent l’utilisation successive d’une classification pixellaire et d’un partitionnement en régions. Par exemple, Huang et al. [Huang *et al.*, 1992] proposent de segmenter grossièrement l’image à l’aide de la méthode de *scale space filtering* sur les histogrammes. Ils affinent ensuite la segmentation par une approche par champs de Markov. Chen et Lu proposent dans [Chen et Lu, 2002] un algorithme de segmentation basé à la fois sur une classification floue dans l’espace colorimétrique et une segmentation basée région au niveau spatial. L’algorithme de classification floue permet de regrouper les pixels proches au sens colorimétrique. Ces classes de pixels sont ensuite fusionnées afin de former des régions de pixels connexes dans l’image. Plusieurs algorithmes basés sur la même idée sont également proposés dans [Chen *et al.*, 2003]. Enfin, dans [Lezoray et Cardot, 2002a] est étudiée la coopération entre classification et algorithme de montée des eaux, la classification servant à déterminer les marqueurs de la montée des eaux. Les mêmes auteurs ont également proposé une approche de segmentation hybride initialisée par une classification d’histogrammes bidimensionnels et achevée par fusion de régions et algorithme de partage des eaux.

3.5 Discussions et conclusions

Beaucoup d'auteurs préconisent souvent un type de méthodes de segmentation par rapport aux autres. Ce qu'il faut prendre en compte dans un premier temps avant de privilégier une technique, c'est bien évidemment l'application concernée, ainsi que le type d'images à traiter. Avant de choisir un type de méthodes de segmentation couleur, on peut se poser entre autres les questions suivantes :

- Admet-on la supervision de l'utilisateur au cours de la segmentation ?
- Comment qualifier l'image à traiter ? Est-ce une image texturée ? Est-elle contrastée ? Montre-t-elle des objets de couleurs très hétérogènes ou pas ?
- Quel espace d'attributs couleur utiliser ?
- Quelles sont les contraintes temporelles posées par l'application ? Privilégions-nous la qualité de la segmentation à la rapidité des algorithmes ?

La détection de contours par exemple, qui consiste en une recherche des discontinuités dans l'image, s'applique plus facilement à des images peu texturées, aux transitions bien marquées. Ces techniques sont faciles à mettre en œuvre et peu coûteuses en temps de calcul, mais elles aboutissent souvent à une mauvaise fermeture des contours et donc à une mauvaise délimitation des régions. Elles peuvent s'appliquer à la détection d'une primitive de type région, si des post-traitements de reconnaissance de forme peuvent être mis en œuvre. Ce n'est pas le cas de notre étude, où nous supposons que l'élément à détecter est de forme inconnue.

Les méthodes de classification de type clustering analysent la répartition des couleurs dans le système de représentation colorimétrique. Elles supposent que les couleurs présentes dans l'image forment des groupements de pixels correspondant aux régions dans l'image. Ces techniques sont généralement simples dans leur mise en œuvre. Par contre, quand elles ne nécessitent pas la connaissance *a priori* du nombre de classes en présence, elles requièrent le choix de multiples paramètres, ce qui s'effectue souvent de manière empirique. L'utilisation des méthodes d'analyse d'histogrammes suppose que la distribution des couleurs forme des pics de densité représentatifs des classes en présence. Cependant, cette propriété est loin d'être vérifiée par n'importe quelle image. L'analyse d'histogrammes est très efficace lorsque les régions présentes dans l'image sont très distinctes au niveau colorimétrique, très homogènes, mais également quand elles ont des tailles comparables. Dans le cas contraire, elle aboutit généralement à une sous-segmentation. Par contre, ce type d'approches ne nécessite pas la connaissance *a priori* du nombre de classes. Dans le cas d'une extraction d'une seule primitive, la classification est intéressante si la primitive à détecter possède des caractéristiques couleur approximativement connues. En outre la région à détecter doit être homogène et bien distincte de toute autre région dans l'image. Dans le cas contraire, il est nécessaire de procéder à un post-traitement d'étiquetage en composantes connexes, ou d'utiliser des attributs autres que la couleur.

Les méthodes de segmentation par approche ascendante (croissance de régions) ou structurale (division-fusion, ou fusion) opèrent directement dans l'espace géométrique, en exploitant des critères de proximité colorimétrique et spatiale. Elles répondent exactement à la définition donnée page 49. L'ajustement des seuils, des paramètres de contrôle de la propagation, et des

prédicats est loin d'être aisée mais cette diversité des réglages permet d'appliquer la croissance de région à tout type d'images. Par contre la séquentialité de cette technique entraîne un coût important en temps de calcul. Elle se justifie largement lorsque la qualité de la segmentation est privilégiée à l'encontre de sa rapidité, ou encore lorsque l'on veut effectuer une partition grossière de l'image, en vue de traitements supplémentaires.

Les méthodes structurales de type division-fusion s'appliquent également à tout type d'images. Par contre elles requièrent l'utilisation de structures de données assez complexes à mettre en œuvre et pèchent par leur fort coût de calcul. Qui plus est, selon la structure de données choisie, elles ne respectent pas forcément la géométrie de l'image ou les contours réels des objets. Puisque la détection de primitive ne doit pas requérir une partition globale de l'image, les algorithmes de division-fusion ne sont pas appropriés à notre cadre d'étude.

De même, les techniques de croissance de régions peuvent être intéressantes dans le cas où tous les germes détectés sont associés à la région à extraire. Aucun attribut supplémentaire relatif à la forme ou à la texture de l'objet n'est alors requis. Dans le cas où l'image est complexe et comprend différentes régions autres que la primitive à détecter, une classification pixellaire peut permettre de sélectionner ce germe. Cela montre tout l'intérêt de l'utilisation d'une approche mixte, alliant classification et segmentation en régions.

Comme nous l'avons précisé auparavant à l'issue du chapitre 2, le choix des attributs couleur est très intimement lié à l'application et au type d'images traité. Dans le cas d'une classification ou dans le cas d'une détection de primitive, la performance dépend étroitement de l'espace de représentation utilisé, et il faut donc s'attacher à choisir un espace d'attributs suffisamment discriminant des classes en présence. Le choix de l'espace dépend généralement de l'algorithme, du type d'images à traiter et de l'application visée. Face à cette problématique, la sélection automatique de l'espace de représentation des données [Meas-Yedid *et al.*, 2004] ou encore la formation automatique d'un espace hybride peut permettre d'optimiser une classification [Vandenbroucke *et al.*, 2003].

Un problème qui reste encore ouvert est celui de l'évaluation de la qualité de la segmentation. Certes, la comparaison des résultats de segmentation avec ceux d'une segmentation experte, effectuée « à la main », est un moyen simple pour évaluer la validité d'une approche de segmentation [Lezoray et Cardot, 2002a]. Pour s'affranchir de l'intervention de l'utilisateur, des auteurs tels que Liu *et al.* [Yang et Liu, 1994] préconisent des critères automatiques d'évaluation de la segmentation, se basant sur des paramètres tels que le nombre et la taille des régions formées, sur le calcul de critère d'homogénéité des régions et d'hétérogénéité entre régions adjacentes. Par la suite, Borsotti *et al.* [Borsotti *et al.*, 1998] améliorèrent ce critère de manière à ce qu'il pénalise plus durement la formation de petites régions et qu'il ne favorise plus les sous-segmentations. Des variantes de ce critère sont exploitées dans [Meas-Yedid *et al.*, 2004], sur des images microscopiques, dans le but de sélectionner automatiquement l'espace couleur. D'après [Trémeau *et al.*, 2004], ce critère correspond de manière assez proche au jugement visuel, mais cela reste encore très dépendant de l'application.

Le chapitre qui suit s'intéresse à la détection d'une primitive couleur transparente (de type marquage à l'encre), de forme inconnue, et pour laquelle les pixels ne sont pas forcément tous

connexes. Nous verrons que l'analyse éclairée des attributs couleur et la mise en œuvre d'une technique de segmentation mixte, alliant à la fois une classification et une croissance de région, permettent de répondre à cette problématique.

DÉTECTION D'UN MARQUAGE À L'ENCRE

Ce chapitre s'intéresse à la détection d'un marquage à l'encre, dont la quantité d'encre et la couleur du support varient d'une image à l'autre. Nous montrons que cette tâche s'apparente à un problème de transparence. L'utilisation de la couleur, étudiée dans le chapitre 2, nous permet de définir des attributs quasiment invariants vis-à-vis de la concentration en encre, de la couleur du fond sur lequel est imprimé le marquage et de la couleur de l'encre. Ces attributs sont ensuite exploités dans le cadre de la segmentation. L'étude du chapitre 3 nous amène quant à elle à choisir un type d'approche de segmentation mixte adaptée à notre cadre d'étude, où le marquage à détecter est de forme inconnue. Plusieurs variantes sont proposées suivant le type d'images. Enfin, quelques résultats de segmentation sont présentés. Par ailleurs, le chapitre 5 se servira largement de l'étude effectuée ici dans le cadre de la traçabilité de produits porcins. Notons que ce chapitre représente une amélioration des travaux décrits dans [Gouiffès et al., 2004a, Gouiffès et al., 2004b].



De nombreuses applications de vision par ordinateur requièrent la détection d'un même objet dans une série d'images couleur similaires. Citons par exemple les applications de détection de route pour la robotique mobile [Lin et Chen, 1991], la séparation plantes/sol pour le machinisme agricole [Foucher *et al.*, 2001], ou encore l'évaluation de la qualité en agro-alimentaire [Raffy, 1999, Legard, 1999].

Ce chapitre se focalise sur la détection d'un marquage à l'encre. Il s'agit de segmenter de manière automatique et répétable ce type de primitive, dont seule la couleur de l'encre est connue *a priori*. Nous supposons à la fois que sa forme est indéterminée et qu'elle ne montre pas de texture discriminante. De plus, les images à traiter peuvent contenir d'autres éléments susceptibles d'introduire des ambiguïtés dans la tâche de détection.

L'état de l'art des méthodes de segmentation couleur a largement insisté sur l'importance du choix des attributs couleur. Or, la couleur d'un marquage au jet d'encre est un paramètre difficile à contrôler. Selon l'opacité et la quantité d'encre utilisée, la couleur du marquage peut

dépendre de la couleur du support sur lequel il est effectué. La variabilité de la reflectance de l'encre est problématique lorsqu'il s'agit de caractériser ou de segmenter cette encre, quels que soient son pouvoir couvrant, sa quantité ou la couleur du support sur lequel elle est imprimée. C'est pourquoi nous allons exploiter les caractéristiques de transparence de l'encre pour en déduire des attributs couleur facilitant au mieux la détection d'un marquage à l'encre, quelle que soit la qualité de l'impression.

Ce chapitre s'articule ainsi. La section 4.1 étudie le problème général du phénomène de transparence et sa modélisation dans l'espace colorimétrique. Nous verrons dans la section 4.2 que la variation de la couleur de l'encre en fonction de sa concentration et de la couleur du fond peut être apparentée au problème plus général de transparence. Nous en déduisons alors des attributs couleur susceptibles de faciliter la détection automatique et répétable d'un marquage à l'encre. La section 4.3 propose ensuite un algorithme de segmentation basé sur l'utilisation des attributs invariants. Enfin, des résultats de segmentation valident les approches proposées.

4.1 La transparence

À notre connaissance, le problème de la détection d'un marquage à l'encre sur un fond uniforme n'a pas été abordé dans la littérature. En revanche, la transparence de filtres colorés a fait l'objet de nombreux travaux. La perception de ce phénomène, qui a concerné de nombreuses études psychosensorielles [Beck *et al.*, 1984, Braddick *et al.*, 2002], n'est possible que lorsque l'objet transparent, un filtre par exemple, recouvre partiellement plusieurs objets de couleurs différentes, tel que le montre la figure 4.1.

La façon la plus simple de modéliser l'effet de transparence est une pure translation, soit de luminance dans le cas des images en niveaux de gris, soit de couleur. La couleur d'un fond a est translatée d'un vecteur t dans l'espace colorimétrique, et le résultat de la superposition est une couleur r telle que $r = a + t$.

Un modèle plus précis, le modèle de convergence de [Metelli, 1970], exprime l'effet de transparence comme un changement de luminance additif. En effet, si l'on considère un filtre transparent de luminance b , posé sur un fond de luminance a , alors la luminance r résultant de la superposition du filtre sur le fond est donnée par $r = \gamma b + (1 - \gamma)a$ où γ représente l'opacité du filtre. Ce modèle a été généralisé au cas tridimensionnel de la couleur dans [Pos, 1989], puis plus récemment amélioré dans [Faul et Ekroll, 2002]. Dans [D'Zmura *et al.*, 1997] le lecteur pourra trouver une description des différents modèles de transparence. Dans le cas tridimensionnel de la couleur, la couleur résultante r de la superposition d'un filtre transparent de couleur b sur un fond de couleur a est donné par : $r = \gamma b + (1 - \gamma)a$. Ainsi, tous les pixels se situant sous un filtre uniforme subissent une translation vers la couleur ce filtre. De plus, le phénomène de transparence apparaît avec des translations équilumineuses sur chaque couleur [D'Zmura *et al.*, 1997].

De nombreux travaux ayant trait à la transparence concernent le domaine de la synthèse d'images. Citons par exemple la modélisation d'un filtre transparent se déplaçant dans une scène [Matsuoka *et al.*, 2002], ou encore la représentation fidèle d'un filtre [D'Zmura *et al.*, 2000] où les auteurs introduisent de surcroît la notion de changement de contraste.

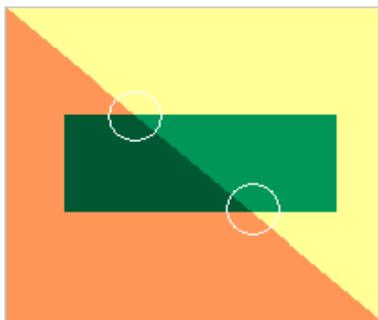


FIG. 4.1 – Filtre vert sur un fond bicolore. Deux jonctions X sont entourées.

En revanche, peu de travaux existent en traitement d'images. Le principal axe de recherche concerne la détection automatique de filtres transparents dans les images couleur. Ceci implique que le filtre transparent recouvre partiellement plusieurs régions de l'image [D'Zmura *et al.*, 1997, Colantoni, 1998]. Il s'agit alors de détecter des *jonctions* X dans l'image, comme le décrit la figure 4.1. Sur cette image, un filtre transparent recouvre partiellement deux régions de couleurs différentes et uniformes. Il se forme alors une jonction X entre quatre régions de couleurs différentes. L'opacité γ peut alors être calculée à partir de deux régions recouvertes par un même filtre.

Toutefois, dans de nombreuses applications, il est impossible d'obtenir de telles informations. Le problème que nous nous posons ici consiste à détecter l'élément transparent b , quelle que soit l'opacité et quelle que soit la couleur du fond a , sans avoir recours aux jonctions X . Pour ce faire, nous effectuons une mesure de l'opacité, que nous mettons en relation avec la concentration de l'encre utilisée. D'autre part, nous définissons un attribut invariant par rapport à la concentration de l'encre. Ces différentes caractéristiques sont exploitées dans le cadre d'une segmentation couleur d'un marquage imprimé à l'encre.

4.2 Étude de la couleur de l'encre imprimée sur un support de couleur uniforme

Comme nous l'avons évoqué précédemment, la couleur d'un marquage imprimé sur un support dépend fortement de la concentration de l'encre et de la couleur de l'objet sur lequel elle est appliquée (voir les figures 4.2a et 4.2b). La figure 4.2a représente une encre noire appliquée sur deux fonds différents de couleurs respectives verte et rose, avec des quantités d'encre relativement variables. Quant à la figure 4.2b, elle correspond à l'impression de trois encres de couleurs différentes sur un support de même couleur. Les figures 4.2c et 4.2d illustrent respectivement les distributions couleur dans l'espace RGB des pixels des images 4.2a et 4.2b. Nous allons voir que, dans ces deux cas de figure, il est possible de déterminer un *locus* des couleurs du marquage.

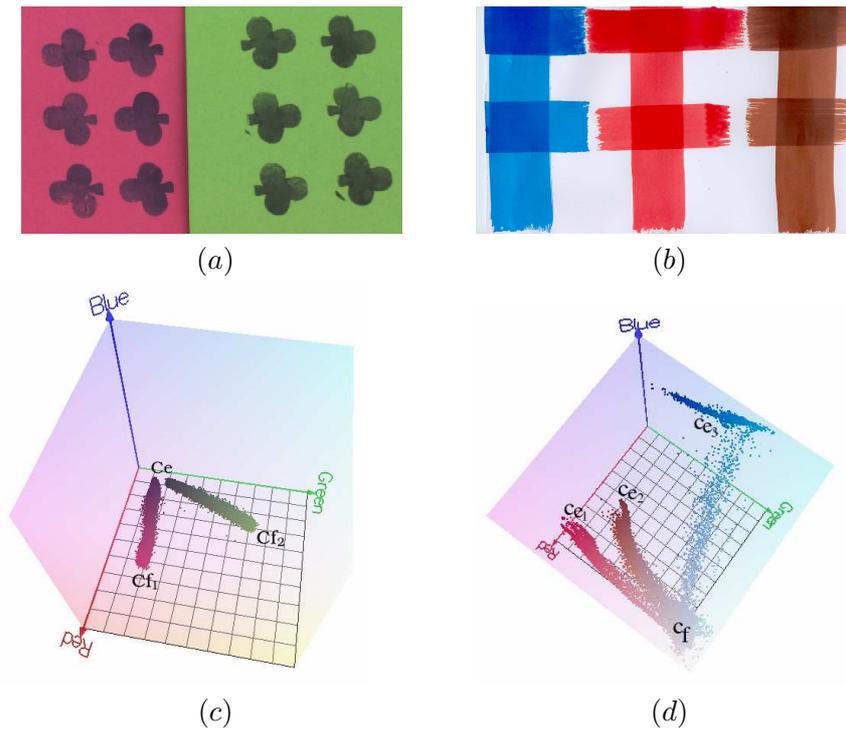


FIG. 4.2 – Étude de la distribution des couleurs de l'encre imprimée sur un fond uniforme. (a) Une même encre noire est imprimée sur deux fonds différents (de couleurs rose et verte) en quantités variables. (b) Trois encres différentes (bleue, rouge et brune) sont imprimées sur un fond blanc. (c) Distribution des couleurs de l'image 4.2a dans l'espace RGB. (d) Distribution des couleurs de l'image 4.2b dans l'espace RGB.

4.2.1 Locus des couleurs associées au marquage à l'encre

Sur les deux figures 4.2c et 4.2d, toutes les couleurs c associées aux marquages sont situées approximativement sur un segment $[C_e, C_f]$, C_f étant la couleur du support et C_e la couleur de l'encre quand il n'y a aucun effet de transparence, c'est-à-dire pour une quantité (ou concentration) maximale d'encre.

Remarque : Ce postulat n'est plus valable lorsque la couleur du fond ou de l'encre sont saturées, du fait de l'acquisition. C'est par exemple le cas de l'encre bleue C_{e_3} sur la figure 4.2d, dont la distribution des couleurs forme un angle en limite de cube RGB .

Par projection du segment $[C_e, C_f]$ sur les sous-espaces dichromatiques (C_i, C_j) , avec i et $j \in \{RGB\}$, et $i \neq j$, comme l'illustre la figure 4.3, on obtient l'équation de trois droites, reliant deux à deux chacune des coordonnées couleur de l'encre : $C_i = a_{ij}C_j + b_{ij}$. a_{ij} et b_{ij} représentent respectivement le coefficient directeur et l'ordonnée à l'origine du segment $[C_e, C_f]$ exprimé dans le sous-espace (C_i, C_j) . Les coefficients directeurs de ces droites s'obtiennent facilement :

$$a_{ij} = (C_{f_i} - C_{e_i}) / (C_{f_j} - C_{e_j}) \quad (4.1)$$

De même, les ordonnées aux origines de ces droites s'expriment comme suit :

$$b_{ij} = (C_{e_i}C_{f_j} - C_{e_j}C_{f_i}) / (C_{f_j} - C_{e_j}). \quad (4.2)$$

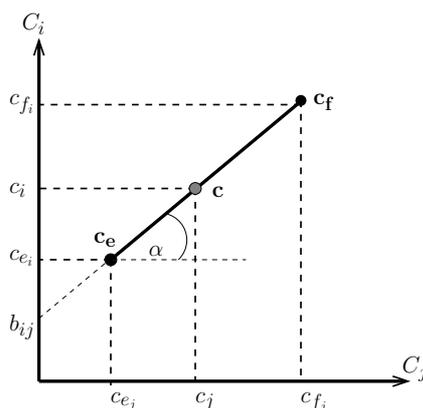


FIG. 4.3 – Projection du segment $[C_e, C_f]$ sur le plan dichromatique (C_i, C_j) .

Dans la suite, nous considérons les trois plans dichromatiques $\{RB\}$, $\{GR\}$ et $\{BG\}$.

Puisque le modèle de transparence proposé n'est qu'une approximation, les couleurs associées à un marquage à l'encre ne se situent pas exactement sur le segment $[C_e, C_f]$. Par contre, l'appartenance d'une couleur $C = (C_R, C_G, C_B)$ à la région $\{support, encre\}$ peut être évaluée en calculant sa distance d_{ij} pour $ij \in \{RB\}, \{GR\}$ et $\{BG\}$ par rapport aux droites de locus de l'encre :

$$d_{ij} = |C_i(p) - (a_{ij}C_j(p) + b_{ij})| \quad (4.3)$$

La distance euclidienne $d_E(c)$ de la couleur C par rapport au segment $[C_e, C_f]$ peut également être utilisée. Plus cette distance est faible et plus la couleur est susceptible d'appartenir à l'ensemble $\{support, encre\}$.

L'utilisation de ces distances n'est valide que si les coefficients a_{ij} et b_{ij} peuvent être obtenus de manière suffisamment précise. Par conséquent, il s'avère primordial d'obtenir C_e et C_f de manière précise. Ce point sera abordé dans la suite, notamment dans le paragraphe 4.3.1.

Les images 4.4a, 4.5a et 4.6a correspondent à un cas concret de problème de transparence de l'encre sur trois supports de couleurs différentes : respectivement vert, jaune et gris. Sur ces supports sont appliquées quatre encres de couleurs différentes : noire bleutée N (zones A, B et C), bleue B (zones D, E et F), rouge R (zones G, H et I) et brune Br (zones J, K et L), utilisées en concentrations variables. Les variations de transparence sont bien visibles. Les images binarisées 4.4b, 4.5b et 4.6b correspondent au locus $\{support, encre noire\}$ (représenté par les pixels blancs dans l'image). Ce locus a été obtenu en seuillant l'image des distances euclidiennes des pixels vis-à-vis du locus du marquage. Quelles que soient la couleur du fond et la concentration d'encre, les marquages à l'encre noire sont quasiment tous détectés. Certes les ambiguïtés ne sont pas toujours écartées. Nous constatons en effet, que quelques pixels associés à la zone J (encre brune) sont assimilés, selon notre critère, à de l'encre noire (voir l'image 4.4b).

Plus généralement, ce critère peut s'avérer limitatif dans le cas précis où il s'agit de différencier deux encres de très faible concentration. Dans ce cas, leurs couleurs respectives se

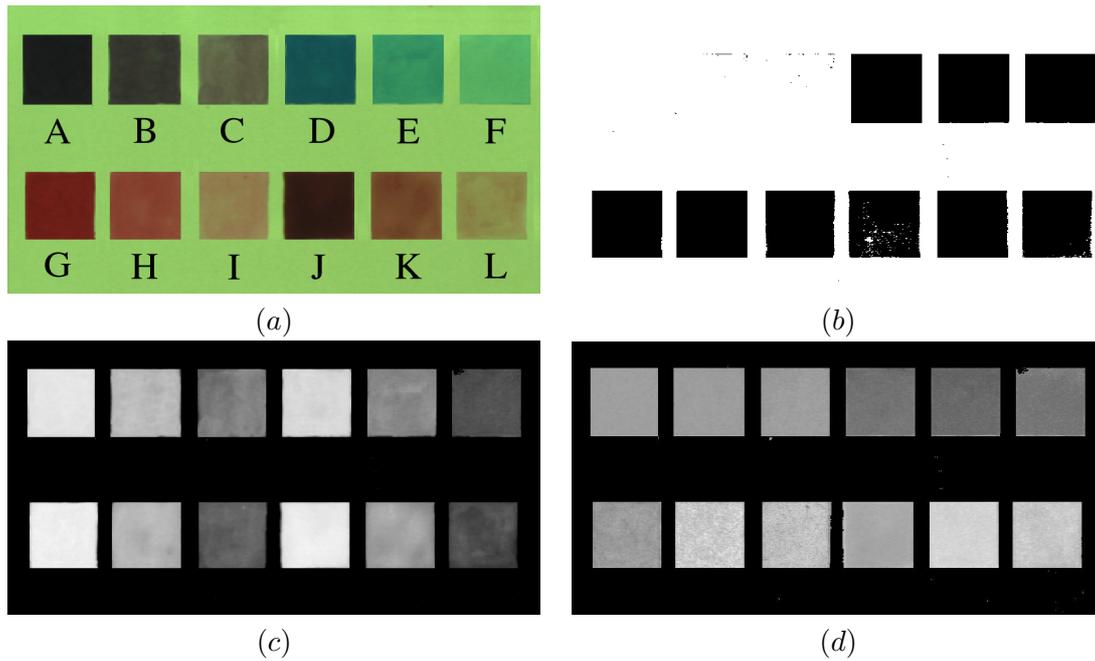


FIG. 4.4 – Attributs invariants de l'encre imprimée sur un support vert. (a) Image RGB de différentes encres de couleur : noire bleutée N (zones A, B, C), bleue B (zones D, E et F), rouge R (zones G, H et I) et brune Br (zones J, K et L). (b) Image binaire représentant la région {support, encre noire} (région en blanc). Elle est obtenue par analyse du critère (1). (c) Image du coefficient de concentration global Q . (d) Exemple de rapport de coefficients de concentration marginaux λ^{RG} .

rapprochent toutes deux de la couleur du fond, et le critère de distance peut évidemment s'avérer insuffisant.

4.2.2 Les coefficients de concentration

Le critère de distance par rapport au *locus*, décrit précédemment, permet en général de déterminer la classe des couleurs associées au marquage à l'encre. Nous introduisons ici de nouveaux attributs appelés *coefficients de concentration*, qui permettent de quantifier la concentration (ou quantité, opacité) de l'encre appliquée sur le support. Ces coefficients peuvent être déterminés marginalement sur chacun des plans couleur, ou globalement dans l'espace RGB . Les *coefficients de concentration marginaux*, appelés Q_{C_i} pour $C_i \in \{RGB\}$ par rapport à la couleur du support C_{f_i} , sont déterminés par la position relative de la couleur sur le segment $[C_{e_i}, C_{f_i}]$:

$$Q_{C_i} = \frac{C_i - C_{f_i}}{C_{e_i} - C_{f_i}} \text{ pour } C_{e_i} \neq C_{f_i} \quad (4.4)$$

Leurs valeurs sont ainsi comprises entre 0 et 1. Effectivement, Q_{C_i} est égal à l'unité, et donc maximal lorsque $C_i = C_{e_i}$, c'est-à-dire lorsque l'encre est au maximum de concentration et sans effet de transparence. À l'inverse, Q_{C_i} est nul, et donc minimum, lorsque $C_i = C_{f_i}$, à savoir lorsque l'encre est à son minimum de concentration. Dans le cas où $C_{e_i} = C_{f_i}$, Q_{C_i} n'est

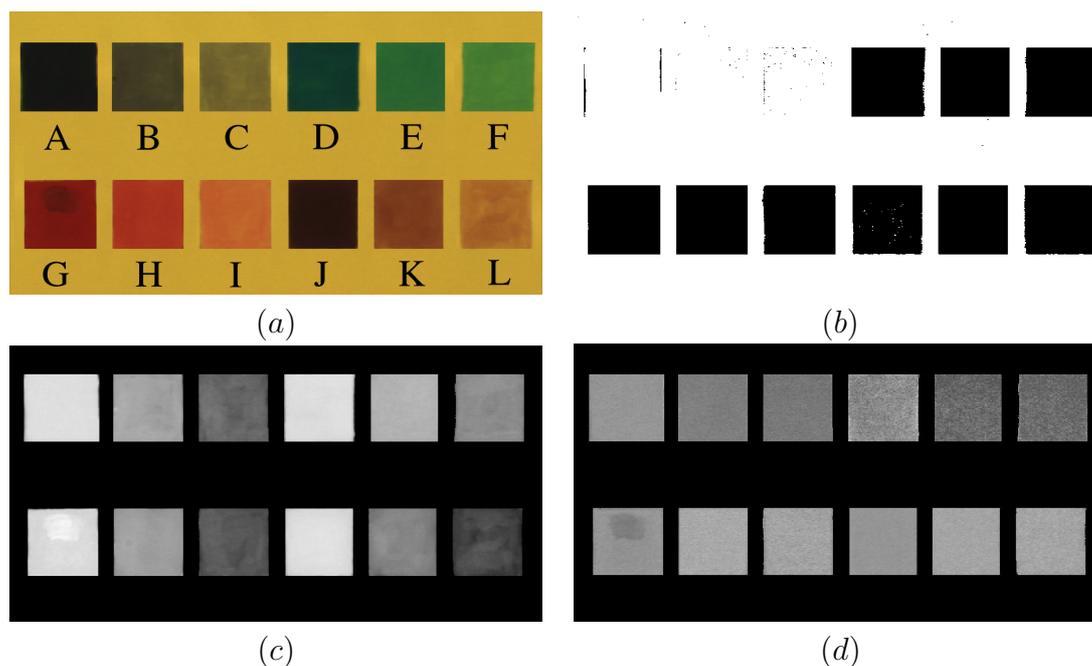


FIG. 4.5 – Attributs invariants de l'encre imprimée sur un support jaune. (a) Image RGB de différentes encres de couleur : noire bleutée N (zones A, B, C), bleue B (zones D, E et F), rouge R (zones G, H et I) et brune R (zones J, K et L). (b) Image binaire représentant la région {support, encre noire} (région en blanc), obtenue par analyse du critère (1). (c) Image du coefficient de concentration global Q . (d) Exemple de rapport de coefficients de concentration marginaux λ^{RG} .

pas défini. Néanmoins, si $C_e \neq C_f$, il existe toujours un autre plan dichromatique (C_i, C_j) de l'espace RGB, pour lequel un autre Q_{C_i} existe.

Définissons également un coefficient de concentration global Q :

$$Q = \frac{\|C - C_f\|}{\|C_e - C_f\|} \text{ pour } C_e \neq C_f \quad (4.5)$$

où C , C_f et C_e correspondent respectivement aux vecteurs couleur du pixel considéré, du support et de l'encre.

Les coefficients de concentration marginaux Q_{C_i} et le coefficient de concentration global Q dépendent uniquement de la concentration de l'encre appliquée sur un fond (c'est-à-dire de la position de la couleur sur le segment $[C_e, C_f]$). Ils sont indépendants de la couleur de l'encre et de la couleur du support sur lequel elle est appliquée.

Si l'on considère la modélisation de la transparence [D'Zmura *et al.*, 1997] :

$$C = \gamma C_e + (1 - \gamma) C_f \quad (4.6)$$

il est aisé de montrer, d'après (4.5), que Q correspond exactement à l'opacité γ .

De même, d'après (4.4), la composante C_i s'exprime à partir des composantes couleur de l'encre et du fond de la manière suivante :

$$C_i = Q_{C_i} C_{e_i} + (1 - Q_{C_i}) C_{f_i} \quad (4.7)$$

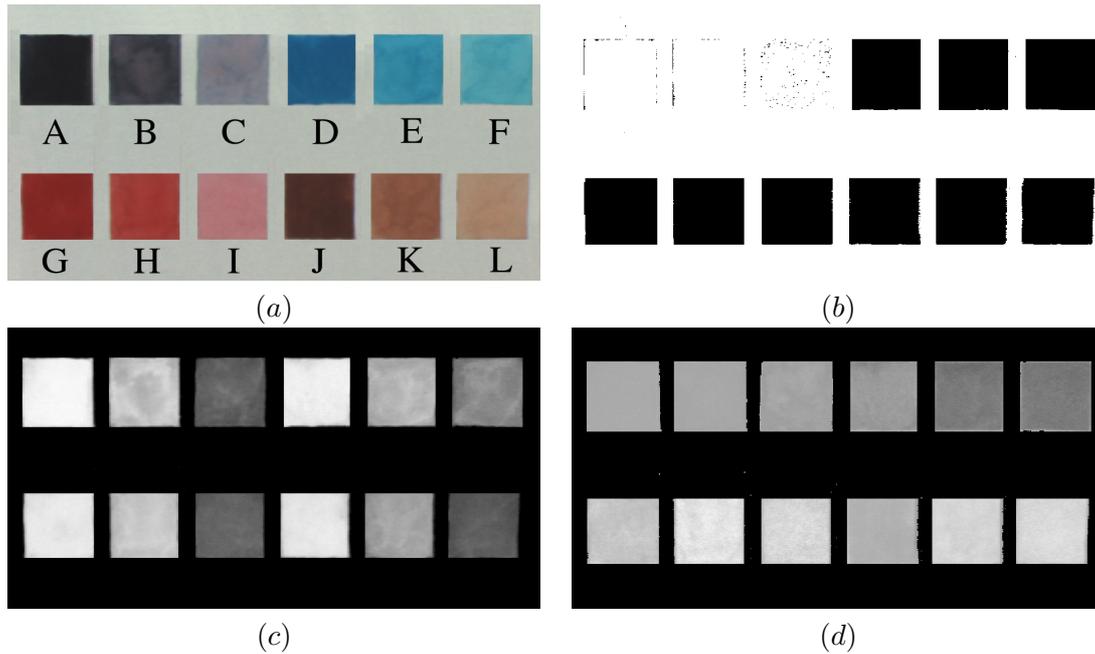


FIG. 4.6 – Attributs invariants de l'encre imprimée sur un support gris. (a) Image RGB de différentes encres de couleur : noire bleutée N (zones A, B, C), bleue B (zones D, E et F), rouge R (zones G, H et I) et brune Br (zones J, K et L). (b) Image binaire représentant la région {support, encre noire} (région en blanc), obtenue par analyse du critère (1). (c) Image du coefficient de concentration global Q . (d) Exemple de rapport de coefficients de concentration marginaux λ^{RG} .

Q_{C_i} n'est autre que l'opacité γ_i exprimée sur la composante C_i . D'après la relation théorique (4.6), on a l'égalité suivante $\gamma_R = \gamma_G = \gamma_B = \gamma$. Or des écarts entre les paramètres γ_R , γ_G , γ_B et le paramètre global γ peuvent être observés, notamment du fait l'approximation de la modélisation et du bruit de mesure. Cette modélisation permet ainsi d'obtenir des coefficients marginaux de la concentration sur chacun des plan RGB .

La modélisation de la transparence au cas d'un filtre peut donc s'appliquer au cas de la transparence de l'encre, c'est-à-dire à la mesure de sa concentration ou de son opacité.

Vérifions ce résultat expérimentalement. Dans un premier temps, considérons les figures 4.4c, 4.5c et 4.6c qui correspondent au coefficient de concentration global Q . Par comparaison de ces images avec les images originales 4.4a, 4.5a et 4.6a, on peut remarquer que les valeurs élevées de Q correspondent effectivement à une forte opacité de l'encre. Pour chacun des trois supports, les marquages à l'encre sont effectués avec des concentrations d'encre comparables (par exemple, la zone A de l'image 4.4a est obtenue avec la même concentration d'encre que la zone A des images 4.6a et 4.5a, etc). Les tableaux 4.1, 4.2 et 4.3 comparent, en fonction de la couleur du support, les valeurs des moyennes RGB et la valeur du coefficient de concentration global Q , obtenues respectivement pour les zones C (encre noire), E (encre bleue) et I (encre rouge). Malgré les variations de couleur d'une zone selon le fond utilisé, les valeurs Q obtenues restent comparables d'un support à l'autre. Par conséquent, le coefficient Q correspond à une mesure corrélée à la concentration, car il dépend peu de la couleur du support. Ajoutons que

l'image des concentrations Q offre une grande discrimination entre pixels de forte concentration et pixels de faible concentration.

TAB. 4.1 – Mesures des moyennes R , G , B et Q obtenues sur la zone C (encre noire).

fond	R	G	B	Q
gris	0.4654	0.4657	0.5009	0.4002
jaune	0.4437	0.3891	0.1768	0.4104
vert	0.3668	0.3811	0.2721	0.4746

TAB. 4.2 – Mesures des moyennes R , G , B et Q obtenues sur la zone E (encre bleue).

fond	R	G	B	Q
gris	0.1984	0.5260	0.6302	0.7087
jaune	0.1636	0.3858	0.1872	0.7005
vert	0.2041	0.5624	0.3711	0.6371

TAB. 4.3 – Mesures des moyennes R , G , B et Q obtenues sur la zone I (encre rouge).

fond	R	G	B	Q
gris	0.6888	0.4565	0.4659	0.4006
jaune	0.7488	0.3778	0.1684	0.3457
vert	0.5670	0.4954	0.3068	0.3789

Dans les cas les plus simples de détection de l'encre sur un fond, l'utilisation du seul coefficient Q peut suffire. Cependant, de manière générale, il est préférable d'exploiter plusieurs coefficients marginaux Q_{C_i} , afin d'augmenter la quantité d'information discriminante. Néanmoins, si l'on considère une image comprenant plusieurs objets, il est possible qu'une région différente de celle du marquage possède des caractéristiques couleur et des coefficients de concentration comparables à la région associée au marquage.

Étant donné que l'approche proposée dépend fortement des valeurs C_{e_i} et C_{f_i} , il est important d'évaluer l'influence d'une erreur de mesure des couleurs de l'encre C_{e_i} et du fond C_{f_i} , sur les valeurs Q_{C_i} . Soient $\sigma_{C_{e_i}}^2$ et $\sigma_{C_{f_i}}^2$ les incertitudes de mesure obtenues sur C_{e_i} et C_{f_i} et soit $\sigma_{Q_{C_i}}^2$ l'incertitude résultante sur l'estimation de Q_{C_i} . On a :

$$\sigma_{Q_{C_i}}^2 = \left(\frac{\partial Q_{C_i}}{\partial C_{e_i}} \right)^2 \sigma_{C_{e_i}}^2 + \left(\frac{\partial Q_{C_i}}{\partial C_{f_i}} \right)^2 \sigma_{C_{f_i}}^2 \quad (4.8)$$

et d'après (4.4) :

$$\frac{\partial Q_{C_i}}{\partial C_{e_i}} = -\frac{Q_{C_i}}{(C_{e_i} - C_{f_i})} \text{ et } \frac{\partial Q_{C_i}}{\partial C_{f_i}} = \frac{Q_{C_i}}{(C_{e_i} - C_{f_i})} - 1 \quad (4.9)$$

Ainsi, plus la couleur de l'encre C_{e_i} est proche de la couleur du fond C_{f_i} et plus les mesures Q_{C_i} sont bruitées. En pratique, il sera préférable d'éliminer au préalable les axes $C_i \in \{RGB\}$ pour lesquels la distance $C_{e_i} - C_{f_i}$ est trop faible.

Introduisons maintenant de nouveaux attributs, les *rappports de coefficients de concentration marginaux*.

4.2.3 Les rapports des coefficients de concentration marginaux

Les rapports de coefficients de concentration, appelés $\lambda^{C_i C_j}$, pour $C_i, C_j \in \{RGB\}$, avec $C_i \neq C_j$, s'expriment de la manière suivante :

$$\lambda^{C_i C_j} = \frac{Q_{C_i}}{Q_{C_j}} \quad \text{pour } Q_{C_j} \neq 0 \quad (4.10)$$

Comme nous l'avons montré dans la section 4.2.2, les paramètres de concentration Q_{C_i} correspondent à l'opacité γ_i . Or, dans le cas idéal où ces paramètres sont tous les trois égaux sur chaque plan couleur, $\gamma_R = \gamma_G = \gamma_B$, le rapport de coefficient de concentration $\lambda^{C_i C_j}$ est égal à l'unité pour tout couple de coordonnées couleur (C_i, C_j) situé exactement sur le segment $[C_e, C_f]$. Par conséquent, les rapports des coefficients de concentration marginaux sont invariants à la concentration de l'encre sur un support de couleur donnée.

Ces coefficients ne sont pas définis lorsque $Q_{C_j} = 0$, c'est-à-dire lorsque $C_i = C_{f_i}$, mais également lorsque l'une des composantes Q_{C_i} ou Q_{C_j} n'est pas définie. Ce cas de figure apparaît lorsque $C_{e_i} = C_{f_i}$ ou $C_{e_j} = C_{f_j}$. Il est alors nécessaire de choisir un autre plan dichromatique (C_i, C_j) . Il existe toujours un plan, pour lequel les coefficients de concentration Q_{C_i} et Q_{C_j} sont définis, pourvu que la couleur de l'identifiant soit différente de celle du support. À titre d'illustration, les figures 4.4d, 4.5d et 4.6d représentent les valeurs du rapport λ^{RG} . Pour une encre donnée, le rapport λ^{RG} est invariant vis-à-vis de la concentration d'encre et de la couleur du support.

Par ailleurs, pour chaque encre $C_e \in \{N, B, R, Br\}$ nous calculons la variabilité \mathcal{V}^{C_e} (en pourcentage) des valeurs $\lambda^{C_i C_j}$ pour les trois plans dichromatiques $\{\{RG\}\{BR\}\{GB\}\}$:

$$\mathcal{V}^{C_e} = \frac{1}{N} \sum_{i,j \in \{RGB\}, i \neq j} \sum_{p \in W_{C_e}} \frac{\lambda^{C_i C_j}(p) - \mu_{\lambda^{C_i C_j}}^{C_e}}{\mu_{\lambda^{C_i C_j}}^{C_e}} \quad (4.11)$$

où $\mu_{\lambda^{C_i C_j}}^{C_e}$ représente la moyenne des valeurs $\lambda^{C_i C_j}$ calculée pour les trois zones ($N = 3$) d'encre de couleur C_e , W_{C_e} est l'ensemble des trois zones de couleur C_e , et p une de ces zones. Les tableaux 4.4, 4.5 et 4.6 décrivent, d'une part la variabilité \mathcal{V}^{C_e} des valeurs λ , obtenues pour chacune des couleurs d'encre C_e , mais également la variabilité associée aux coordonnées RGB .

Dans tous les cas, les rapports $\lambda^{C_i C_j}$ varient moins d'une zone à l'autre, en comparaison par rapport aux simples coordonnées RGB . Là où les coefficients de concentration ne permettent pas d'éliminer toute ambiguïté entre deux couleurs proches, les rapports de coefficients de concentration permettent de caractériser l'appartenance de la couleur au segment $[C_e, C_f]$. Les attributs que nous venons de décrire sont à utiliser avec précaution, étant donné que les valeurs λ sont bruitées dans le cas où l'écart entre C_{e_i} et C_{f_i} est trop faible sur l'une des composantes i .

Dans cette section, nous avons introduit des attributs permettant d'assurer une constance colorimétrique de l'encre vis-à-vis de sa concentration, mais également vis-à-vis du fond sur

TAB. 4.4 – Comparaison entre la variabilité obtenue sur λ et celle obtenue sur les coordonnées RGB, pour les quatre encres imprimées sur le fond vert (figure 4.4).

	\mathcal{V}^N	\mathcal{V}^B	\mathcal{V}^{Br}	\mathcal{V}^R
RGB	37.2	24.6	31.7	41.8
λ	0.4	4.8	9.5	8.8

TAB. 4.5 – Comparaison entre la variabilité obtenue sur λ et celle obtenue sur les coordonnées RGB, pour les quatre encres imprimées sur le fond jaune (figure 4.5).

	\mathcal{V}^N	\mathcal{V}^B	\mathcal{V}^{Br}	\mathcal{V}^R
RGB	38.7	25.0	32.7	38.3
λ	8.7	8.9	7.4	4.2

TAB. 4.6 – Comparaison entre la variabilité obtenue sur λ et celle obtenue sur les coordonnées RGB, pour les quatre encres imprimées sur le fond gris (figure 4.6).

	\mathcal{V}^N	\mathcal{V}^B	\mathcal{V}^{Br}	\mathcal{V}^R
RGB	42.1	20.8	33.0	31.8
λ	1.8	10.0	4.6	7.7

lequel elle est imprimée. Ces attributs sont exploités dans le cadre d'une application de segmentation d'un marquage au jet d'encre.

4.3 Algorithme de segmentation d'un marquage à l'encre

Le choix de l'approche de segmentation peut être guidé en répondant aux questions de la page 68 posées en conclusion du chapitre précédent.

- La détection devant être automatique, toutes les approches qui nécessitent l'intervention de l'utilisateur sont à bannir. Puisque le nombre de classes présentes peut varier d'une image à l'autre, la plupart des techniques de type clustering ne sont pas non plus utilisables.

- Puisque la forme de la région à détecter est inconnue, plusieurs types d'approches peuvent d'ores et déjà être écartés, tels que la détection de contours ou les approches de division-fusion.

- Les contraintes de temps de calcul nous amènent à écarter les méthodes requérant une haute complexité algorithmique, telles que les algorithmes de division-fusion. Du fait que l'image ne comprend pas seulement le marquage à détecter mais bien d'autres éléments pouvant être différents d'une image à l'autre, des techniques de segmentation mixtes peuvent s'avérer tout à fait adéquates.

À notre sens, le type d'approche le plus adapté s'apparente aux méthodes mixtes, alliant à la fois une classification pixellaire et une segmentation dans l'espace géométrique de l'image, en l'occurrence une croissance de région. Les attributs invariants décrits précédemment doivent

permettre la détection d'un marquage dont la quantité d'encre utilisée n'est pas homogène au sein d'une même image, et permettre une bonne robustesse par rapport à une variation de couleur de fond et de la quantité d'encre d'une image à l'autre.

Ainsi, une classification pixellaire à partir de ces attributs doit permettre de sélectionner un premier ensemble de pixels, dits pixels *candidats*. Puisque cette étape n'est pas toujours suffisante pour détecter la globalité du marquage, il est parfois nécessaire d'opérer une étape supplémentaire visant à calculer un critère d'appartenance du pixel au marquage, en fonction du voisinage de celui-ci. La région correspondant au marquage est finalement obtenue après sélection d'un germe, puis croissance de la région associée à ce germe. Préalablement à cette dernière phase, un filtrage spatial de l'image, visant à assurer la connexité des pixels appartenant au marquage, peut être requis.

4.3.1 Remarques préliminaires

Les attributs invariants décrits dans la section 4.2 nécessitent la connaissance de la couleur de l'encre C_e , mais également de la couleur du support C_f . Dans notre cas, nous supposons C_e connue. Elle peut s'obtenir facilement par acquisition d'un *témoin*, c'est-à-dire une zone sur laquelle est appliquée suffisamment d'encre pour éviter toute transparence.

La couleur du support C_f est déterminée par analyse d'histogrammes en supposant que le fond est quasiment uniforme et représente la classe dominante dans l'image. Il est alors représenté par le mode le plus significatif dans les histogrammes couleur RGB . Une solution consiste à simplement assimiler ce mode à une gaussienne de sorte que le maximum de l'histogramme corresponde directement à la valeur maximale de la gaussienne (sa moyenne). Toutefois, une analyse plus approfondie de l'histogramme apporte une meilleure précision. À cet effet, nous proposons l'analyse successive des trois plans couleur RGB , telle qu'elle est expliquée ci-après :

1. Pour chaque composante couleur $C_i = R, G, B$:
 - (a) Calcul de l'histogramme h_i de la composante C_i .
 - (b) Détection de la composante couleur $C_{i_{MAX}}$ majoritaire dans l'histogramme $h_i(C_{i_{MAX}})$.
 - (c) Détection des limites inférieure $C_{i_{LEFT}} < C_{i_{MAX}}$ et supérieure $C_{i_{RIGHT}} > C_{i_{MAX}}$ de la classe du fond, telles que $h_i(C_{i_{LEFT}}) = h_i(C_{i_{RIGHT}}) = \mathcal{S}_h h_i(C_{i_{MAX}})$, avec \mathcal{S}_h un pourcentage fixé.
 - (d) Seuillage du plan couleur i et obtention d'un masque m_i .
2. ET logique entre les trois masques m_i . Obtention d'un masque global m .
3. Calcul de la moyenne colorimétrique de l'image masquée par m .

La figure 4.7 montre les trois histogrammes (figures 4.7a, 4.7b et 4.7c), calculés à partir de l'image 4.6a, ainsi que le fond extrait en prenant $\mathcal{S}_h = 5\%$. Après calcul de la moyenne couleur sur ce fond, on obtient finalement $C_f = (184, 189, 182)$.

Retournons à présent aux coefficients Q_{C_i} et $\lambda^{C_i C_j}$. Comme nous l'avons précisé auparavant, certains attributs marginaux Q_{C_i} ou $\lambda^{C_i C_j}$ ne sont pas définis lorsque l'une des composantes couleur de l'encre est identique à celle du fond. En outre, dans le cas où les deux valeurs sont très proches, les attributs sont susceptibles d'être bruités. C'est pourquoi il est nécessaire de vérifier que ce cas de figure n'intervient pas. Le cas échéant, il faut sélectionner les deux axes

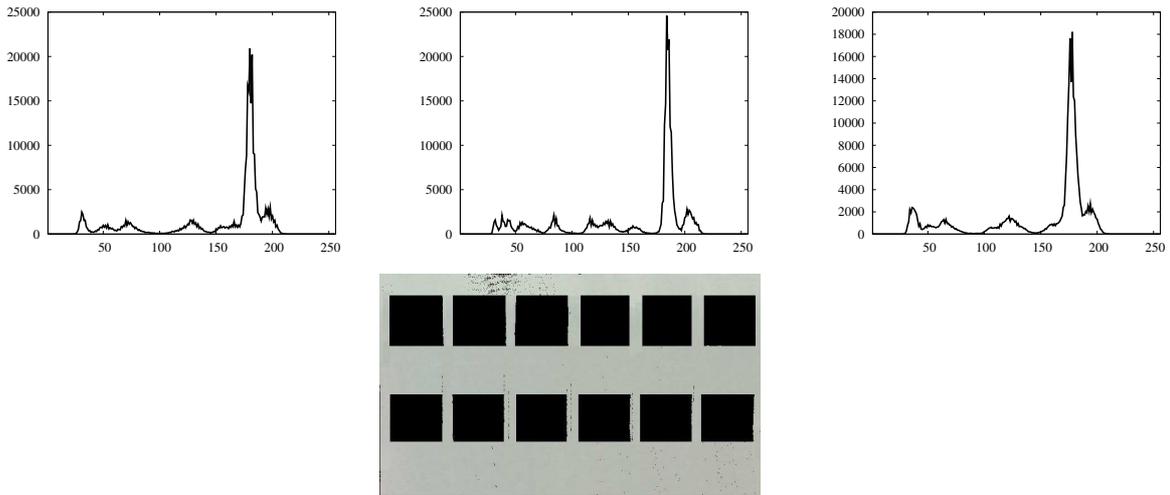


FIG. 4.7 – Sélection du fond pour le calcul de sa couleur.

chromatiques C_i et C_j obtenant la plus grande distance couleur entre C_{e_i} et C_{f_i} d'une part, entre C_{e_j} et C_{f_j} d'autre part. Une alternative consiste à n'utiliser que le coefficient de concentration global.

Les attributs Q_{C_i} ou Q , et $\lambda^{C_i C_j}$ étant calculés, la première étape du traitement de segmentation du marquage procède à la sélection de pixels candidats, c'est-à-dire aux pixels susceptibles d'appartenir au marquage, dans le but d'obtenir un germe pour l'algorithme de croissance de région. Comme la section 4.3.2 s'attache à le montrer, les pixels candidats sont choisis aux endroits de forte concentration d'encre, c'est-à-dire pour des valeurs de coefficient de concentration proches de 1.

4.3.2 Sélection des pixels candidats

Appelons $C_i(p) \in \{RGB\}$ une coordonnée couleur du pixel au point p . Nous proposons ici deux manières de sélectionner les pixels candidats, soit par multi-seuillage soit par calcul de critères plus souples.

4.3.2.1 Par multi-seuillage

Puisque la couleur de l'encre imprimée est localisée sur un segment $[C_e, C_f]$, p appartient au marquage s'il vérifie le critère $\mathcal{C}_1(p)$ suivant [Gouiffès *et al.*, 2004a, Gouiffès *et al.*, 2004b] :

$$\mathcal{C}_1(p) = \begin{cases} 1 & \text{si } \forall (C_i(p), C_j(p)) \in \{RGB\} : |C_i(p) - (a_{ij}C_j(p) + b_{ij})| < \mathcal{S}_d, \\ 0 & \text{sinon} \end{cases} \quad (4.12)$$

S_d est la distance maximale entre un pixel et la droite associée à $[C_e, C_f]$, qui peut être identique sur chacun des plans chromatiques. L'appartenance d'un pixel au *locus* du marquage peut également être caractérisée par le critère qui suit :

$$\mathcal{C}_2(p) = \begin{cases} 1 & \text{si } |\lambda^{C_i C_j}(p) - 1| < S_\lambda \\ 0 & \text{sinon} \end{cases} \quad (4.13)$$

Qui plus est, l'encre possède un pouvoir couvrant donné, c'est-à-dire qu'elle laisse plus ou moins transparaître le support sur lequel elle est imprimée. Cette constatation nous permet de restreindre l'espace d'analyse des Q_{C_i} à un intervalle de valeurs : un pixel p appartient au marquage si sa couleur satisfait un deuxième critère $\mathcal{C}_3(p)$, décrit ci-dessous :

$$\mathcal{C}_3(p) = \begin{cases} 1 & \text{si } \forall C_i(p) \in \{RGB\}, 1 \geq Q_{C_i}(p) > S_{Q_i} \\ 0 & \text{sinon} \end{cases} \quad (4.14)$$

En résumé, un pixel est jugé candidat s'il vérifie le critère suivant :

$$\mathcal{D}_{Cand_1} = \begin{cases} 1 & \text{si } \mathcal{C}_1(p) = \mathcal{C}_2(p) = \mathcal{C}_3(p) = 1 \\ 0 & \text{si } \mathcal{C}_1(p) = 0 \parallel \mathcal{C}_2(p) = 0 \parallel \mathcal{C}_3(p) = 0 \end{cases} \quad (4.15)$$

4.3.2.2 Seconde technique : calcul d'un critère

Afin de réduire le nombre de seuils fixes mis en jeu dans la sélection des candidats par multi-seuillage, on peut préférer l'utilisation de critères plus tolérants. Il faut alors définir des fonctions croissantes par rapport à l'appartenance au marquage. Nous avons choisi de considérer des fonctions gaussiennes \mathcal{D}_Q , \mathcal{D}_λ et \mathcal{D}_d , réglées par leur moyenne et leurs écarts-type non-nuls σ_Q , σ_L et σ_D :

$$\mathcal{D}_Q(p) = \sum_{ij \in RGB} \frac{e^{-(Q_i(p) - \mu_Q)^2}}{\sigma_Q^2} \text{ pour } Q_{C_i} \leq 1, 0 \text{ sinon} \quad (4.16)$$

$$\mathcal{D}_\lambda(p) = \sum_{ij \in RGB} \frac{e^{-(\lambda^{C_i C_j}(p) - 1)^2}}{\sigma_L^2} \quad (4.17)$$

$$\mathcal{D}_d(p) = \sum_{ij \in RGB} \frac{e^{-d_{ij}^2(p)}}{\sigma_D^2} \quad (4.18)$$

Le paramètre μ_Q est la concentration de l'encre recherchée, que l'on fixe typiquement à 1. À partir de ces attributs, un critère global \mathcal{D}_1 , exprimant la potentialité du pixel à être candidat au marquage, peut être obtenu, soit par multiplication soit par combinaison linéaire (de coefficients a , b et c) des critères :

$$\mathcal{D}_1(p) = a\mathcal{D}_Q(p) + b\mathcal{D}_\lambda(p) + c\mathcal{D}_d(p) \quad (4.19)$$

Les coefficients a , b et c permettent de fixer l'importance de chacun des critères. Dans le cas typique, ils sont choisis égaux, de manière à ne privilégier aucun des critères. Le critère par multiplication est plus strict que le critère par combinaison linéaire, puisqu'il suffit qu'un des critères soit nul pour que le critère global s'annule, contrairement à la combinaison linéaire. Par

analyse de ce critère global, un seul seuil \mathcal{S}_{cand} permet de sélectionner les pixels candidats. Un pixel est alors considéré comme candidat lorsqu'il vérifie le critère suivant :

$$\mathcal{D}_{Cand_2}(p) = \begin{cases} 1 & \text{si } \mathcal{D}_1 > \mathcal{S}_{cand}, \\ 0 & \text{sinon} \end{cases} \quad (4.20)$$

En guise d'exemple, les figures 4.9a, 4.9b et 4.9c montrent respectivement les critères $\mathcal{D}_\lambda(p)$, $\mathcal{D}_Q(p)$ et $\mathcal{D}_d(p)$, en prenant comme valeurs de paramètres $\sigma_L=0.2$, $\sigma_Q=0.7$ et $\sigma_D=0.2$. Enfin, le critère $\mathcal{D}_1(p)$, représenté sur la figure 4.9d est calculé avec $a = b = c = 1/3$, pour ne privilégier aucun des critères par rapport aux autres.

Appelons p_s un pixel *candidat*, vérifiant soit le critère \mathcal{D}_{Cand_1} soit le critère \mathcal{D}_{Cand_2} . La sélection des candidats peut autoriser l'apparition de bruit, puisque les pixels associés à d'autres régions de l'image peuvent posséder des caractéristiques colorimétriques parfois très similaires à celles du marquage. Ainsi, l'analyse du voisinage des pixels candidats permet de calculer des scores d'appartenance au marquage, et ainsi éliminer les mauvais pixels candidats.

4.3.3 Élimination des faux candidats par calcul d'un score d'appartenance au marquage

Le *score d'appartenance* au marquage traduit la probabilité qu'un pixel candidat fasse effectivement partie du marquage. Dans certains cas de figure, l'analyse du critère \mathcal{D}_1 , décrit par (4.19), peut suffire à décrire l'appartenance du pixel à la région associée au marquage. C'est le cas des images pour lesquelles le marquage à détecter s'avère très dissemblable de tout autre pixel présent dans l'image. L'image des scores d'appartenance $I_{\mathcal{A}}$ est alors directement donnée par :

$$I_{\mathcal{A}}(p) = \mathcal{D}_1(p) \quad (4.21)$$

Néanmoins, lorsque l'image comprend des éléments très proches colorimétriquement du marquage à détecter, certains pixels de l'image n'appartenant pas au marquage peuvent détenir une valeur \mathcal{D}_1 élevée. Il est alors judicieux de favoriser les groupes de pixels spatialement proches et de valeurs d'appartenance élevées, et d'éliminer les pixels de valeurs élevées mais montrant peu de connexités avec d'autres pixels de forte valeur. Dans ce cas, l'image des scores d'appartenance $I_{\mathcal{A}}$ est obtenue par analyse du voisinage des pixels candidats. Celle-ci pourrait être opérée directement sur l'image initiale, sans une sélection préalable des *pixels candidats*. Toutefois cela augmenterait énormément le coût calculatoire de l'algorithme, puisqu'au lieu d'analyser le voisinage en quelques pixels, cela nécessiterait d'analyser le voisinage de l'ensemble des pixels de l'image.

L'obtention de l'image $I_{\mathcal{A}}$ se base sur le postulat suivant : si un pixel candidat possède plusieurs voisins susceptibles d'appartenir au marquage, alors il a de plus fortes probabilités pour effectivement faire partie du marquage. À l'inverse, si ce candidat ne possède aucun pixel voisin susceptible d'appartenir au marquage, il est fort probable qu'il corresponde à du bruit et dans ce cas il doit être éliminé. Ainsi, pour chaque candidat p_s , un score d'appartenance $I_{\mathcal{A}}(p_s)$ est calculé, en fonction des attributs de son voisinage. Plus il possède de voisins proches au sens colorimétrique, plus le score d'appartenance est élevé. Dans cette hypothèse, les pixels voisins

possèdent approximativement les mêmes coefficients $\lambda^{C_i C_j}$. Là encore deux techniques peuvent être envisagées, par multi-seuillage ou par analyse d'un critère.

4.3.3.1 Multi-seuillage

Pour chaque pixel p appartenant au voisinage V_s d'un pixel p_s , on analyse un critère d'homogénéité $\mathcal{C}_4(p)$, basé sur la distance euclidienne $d_\lambda(p)$ entre les valeurs λ du pixel p_s et son voisin p :

$$\mathcal{C}_4(p) = \begin{cases} 1 & \text{si, } \forall(p) \in V_s, d_\lambda(p) < S_2 \\ 0 & \text{sinon} \end{cases} \quad (4.22)$$

Les pixels voisins doivent être suffisamment proches du *locus* $[C_e, C_f]$ et détenir un coefficient de concentration suffisamment important. Ainsi, ils doivent satisfaire les critères \mathcal{C}'_1 et \mathcal{C}'_3 , similaires aux critères \mathcal{C}_1 et \mathcal{C}_3 avec des seuils $\mathcal{S}_{d_{ij}}$ et \mathcal{S}_{Q_i} moins stricts. Cela permet de prendre en compte des pixels qui appartiennent à l'identifiant bien que leur concentration d'encre soit trop faible pour qu'ils soient sélectionnés initialement comme candidats.

Les pixels pour lesquels aucun voisin ne répond au critère sont éliminés et leur score d'appartenance est annulé. Chaque fois qu'un voisin $p \in V_s$ répond à l'ensemble des critères \mathcal{C}'_1 , \mathcal{C}'_3 et \mathcal{C}_4 , son score d'appartenance ainsi que celui du pixel central augmentent. L'algorithme est énoncé ci-après :

Pour chaque pixel $p \in V_s$:

Si $(\mathcal{C}'_1(p), \mathcal{C}'_3(p), \mathcal{C}_4(p)) = (1, 1, 1)$ alors :

$$\begin{cases} I_A(p_s) & = & I_A(p_s) + 2 \\ I_A(p) & = & I_A(p) + 1 \end{cases} \quad (4.23)$$

Sinon RIEN.

Si, pour tous les pixels $p \in V_s$:

$(\mathcal{C}'_1(p), \mathcal{C}'_3(p), \mathcal{C}_4(p)) = (0, 0, 0)$ alors :

$$\begin{cases} I_A(p_s) & = & 0 \\ I_A(p) & = & 0 \quad \forall(p) \in V_s \end{cases} \quad (4.24)$$

Sinon RIEN.

Cette approche a été développée dans [Gouiffès *et al.*, 2004a, Gouiffès *et al.*, 2004b].

4.3.3.2 Analyse d'un critère \mathcal{D}_{app}

Afin de s'affranchir de la définition de plusieurs seuils fixes, un seul critère d'appartenance \mathcal{D}_{app} peut être défini, de la même forme que (4.19). Plus le critère \mathcal{D}_{app} est élevé, plus le pixel a de fortes probabilités d'appartenir au marquage.

Les pixels courants pour lesquels aucun voisin ne répond au critère sont éliminés et leur score d'appartenance est annulé. Chaque fois qu'un voisin $p \in V_s$ répond aux critères d'appartenance, son score et celui du pixel central augmentent. L'algorithme devient alors :

Pour chaque pixel $p \in V_s$:

$$\begin{cases} I_{\mathcal{A}}(p_s) &= I_{\mathcal{A}}(p_s) + \mathcal{D}_{app}(p) \\ I_{\mathcal{A}}(p) &= I_{\mathcal{A}}(p) + \frac{\mathcal{D}_{app}(p_s)}{\|p_s - p\|} \end{cases} \quad (4.25)$$

$\|p_s - p\|$ représente la distance euclidienne, en pixels, entre le pixel voisin et le pixel central, de manière à accorder une plus grande importance aux pixels les plus proches des pixels candidats. Le résultat de ce traitement est une image en niveaux de gris $I_{\mathcal{A}}(p)$ et dont les valeurs correspondent aux scores d'appartenance calculés. Dans le cas où il n'y a qu'un marquage à repérer, il est intéressant de diviser les valeurs des pixels de l'image $I_{\mathcal{A}}(p)$ par leur distance euclidienne calculée par rapport au centre d'inertie de l'image $I_{\mathcal{A}}(p)$. Cela fait croître le score d'appartenance des pixels appartenant au plus important groupe de pixels de fort degré d'appartenance. De la même manière, cela fait décroître les scores d'appartenance des pixels isolés. La figure 4.9e montre un exemple d'image de critère d'appartenance obtenue.

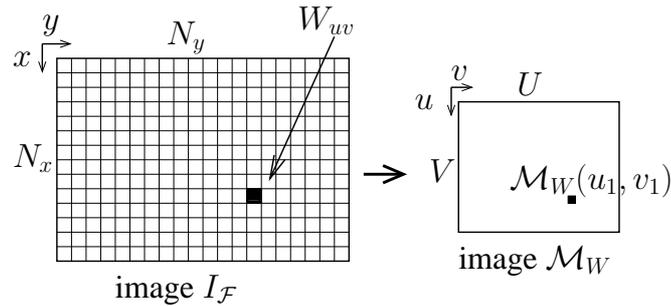
À ce niveau des traitements, la sélection du pixel obtenant le score d'appartenance le plus élevé peut être sélectionné comme un pixel germe pour une croissance de région. Néanmoins, les pixels qui possèdent un score d'appartenance élevé ne sont pas forcément tous connexes. La prochaine section explique comment remédier à ce problème et obtenir une unique composante connexe englobant la totalité du marquage.

4.3.4 Sélection d'une composante connexe associée au marquage

Considérons l'image des scores d'appartenance $I_{\mathcal{A}}$, de taille $N_x \times N_y$. Elle est partitionnée en fenêtres W_{uv} , de taille $M_x \times M_y$. Chaque pixel de $I_{\mathcal{A}}$ appartenant à W_{uv} est remplacé par la moyenne des pixels $I_{\mathcal{A}}$ sur cette fenêtre. Soit $I_{\mathcal{F}}$ cette nouvelle image, qui est schématisée par la figure 4.8.

Afin d'accélérer les traitements qui suivent, une image \mathcal{M}_W de taille $U \times V$ (où $U = \frac{N_x}{M_x}$ et $V = \frac{N_y}{M_y}$) est formée à partir des valeurs des moyennes $I_{\mathcal{F}}$. Chaque pixel $\mathcal{M}_W(u, v)$ contient la valeur moyenne de la fenêtre W_{uv} dans l'image $I_{\mathcal{F}}$ (voir figure 4.8).

L'image $\mathcal{M}_W(u, v)$ nous permet d'extraire un germe du marquage, de manière à effectuer une croissance de région, dont le principe a été exposé dans la section 3.3.1. En effet, le pixel de $\mathcal{M}_W(u, v)$ possédant la valeur maximale, $\mathcal{M}_{W_{MAX}}$ détient la plus forte probabilité d'appartenir au marquage. Il est donc sélectionné comme un germe associé à la composante connexe du marquage. Ensuite, l'algorithme de croissance fait évoluer la région associée au germe tant que le pixel à traiter est supérieur à un pourcentage \mathcal{S}_{croiss} de la valeur $\mathcal{M}_{W_{MAX}}$.


 FIG. 4.8 – Passage de l'image $I_{\mathcal{F}}$ à l'image des moyennes \mathcal{M}_W .

La sélection du germe s'effectue sur l'image des scores d'appartenance. Quant au processus de croissance, il peut être effectué sur l'image associée aux critères de distance, dans lesquelles on aura préalablement éliminé les pixels de concentration trop faible associés au fond.

L'intérêt de cette méthode de filtrage est d'obtenir un masque englobant l'ensemble de pixels *spatialement et colorimétriquement proches mais pas forcément tous connexes*. Par contre, le choix de la taille de la fenêtre dépend de l'image. Elle doit être suffisamment grande de manière à assurer la formation d'une composante connexe de fenêtres de valeurs non nulles permettant de couvrir la série de caractères.

La région formée par croissance de région dans l'image $\mathcal{M}_W(u, v)$ est retransmise dans l'image initiale, de taille $N_x \times N_y$ et forme une composante connexe englobant la totalité du marquage. Les figures 4.9g et 4.9h illustrent l'étape de sélection de la composante connexe associée au marquage, la première figure correspond à l'image après filtrage spatial $I_{\mathcal{F}}$ (avec $M_x = M_y = 8$) et la deuxième au masque finalement obtenu. Notons également qu'il suffit de modifier μ_Q (voir l'équation (4.17)) pour sélectionner un marquage effectué avec la même encre mais avec une concentration différente. C'est ce qu'illustrent les figures 4.9i et 4.9k où les valeurs de μ_Q utilisées sont respectivement égales à 0,6 et 0,4. Les figures 4.9j et 4.9l correspondent quant à elles aux masques obtenus.

Finalement, nous avons défini quatre façons de segmenter un marquage à l'encre à partir des attributs définis dans la section 4.2. La figure 4.10 résume ces quatre variantes, selon qu'elles utilisent une sélection de pixels candidats (méthodes de type A et B) ou non (méthodes de type C et D), un filtrage spatial (B et C) ou non (A et D). La sélection de pixels candidats suivie du calcul d'un score d'appartenance se justifie lorsque l'image comprend des pixels qui ne font pas partie du marquage mais qui possèdent des caractéristiques couleur très similaires à celui-ci. L'analyse de voisinage permet alors de privilégier les régions homogènes au détriment des pixels isolés. Le filtrage spatial est requis lorsque les pixels associés au marquage à détecter ne sont pas tous connexes. C'est le cas des marquages constitués de différents caractères par exemple. Il permet ainsi d'obtenir un masque englobant la totalité du marquage.

Notons que, suivant l'application visée et le type d'image, le nombre d'attributs utilisés peut être réduit pour n'en garder que les plus pertinents. Par exemple, le coefficient de concentration global peut être exploité. On peut également sélectionner le plan dichromatique pour lequel l'écart entre la couleur du fond et la couleur de l'encre est maximal, comme nous l'avons expli-

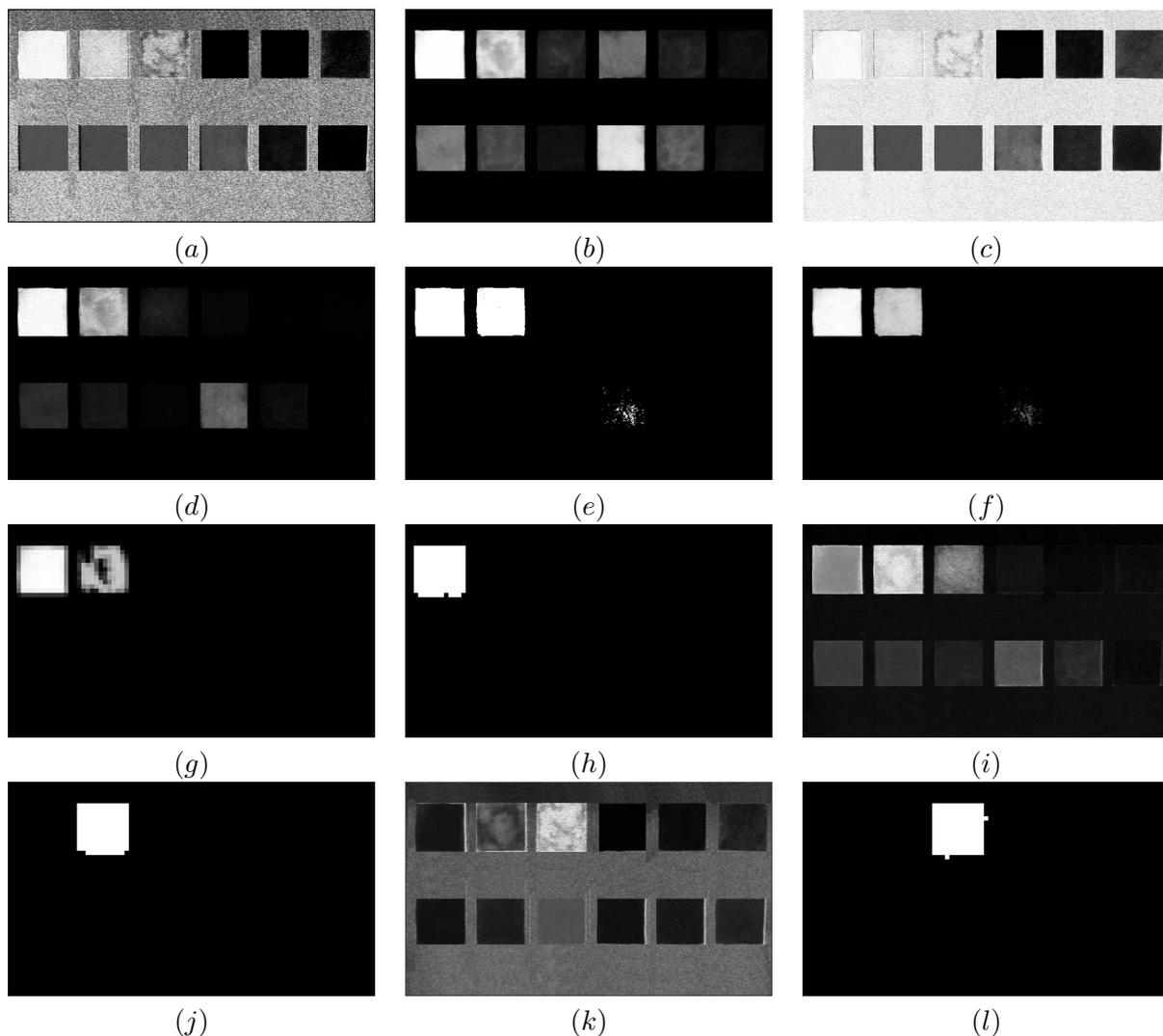


FIG. 4.9 – Illustration de la méthode de segmentation d'un marquage à l'encre, appliquée sur l'image 4.6a. (a) Critère \mathcal{D}_L . (b) Critère \mathcal{D}_Q . (c) Critère \mathcal{D}_d . (d) Critère \mathcal{D}_1 . (e) Image des pixels candidats (critère \mathcal{D}_{Cand_2}). (f) Image des scores d'appartenance I_A . (g) Image après filtrage spatial $I_{\mathcal{F}}$. (h) Masque associé au marquage. (i) Critère \mathcal{D}_1 calculé avec $\mu_Q=0.6$. (j) Masque sélectionné avec $\mu_Q=0.6$. (k) Critère \mathcal{D}_1 calculé avec $\mu_Q=0.4$. (l) Masque sélectionné avec $\mu_Q=0.4$.

qué dans la section 4.3.1. La composante connexe extraite peut être utilisée comme un masque du marquage, sur lequel sont effectués des traitements de segmentation supplémentaires. Elle peut également servir de zone d'intérêt pour une tâche de vision active ou de suivi. Les approches développées sont illustrées dans la prochaine section à l'aide de quelques exemples.

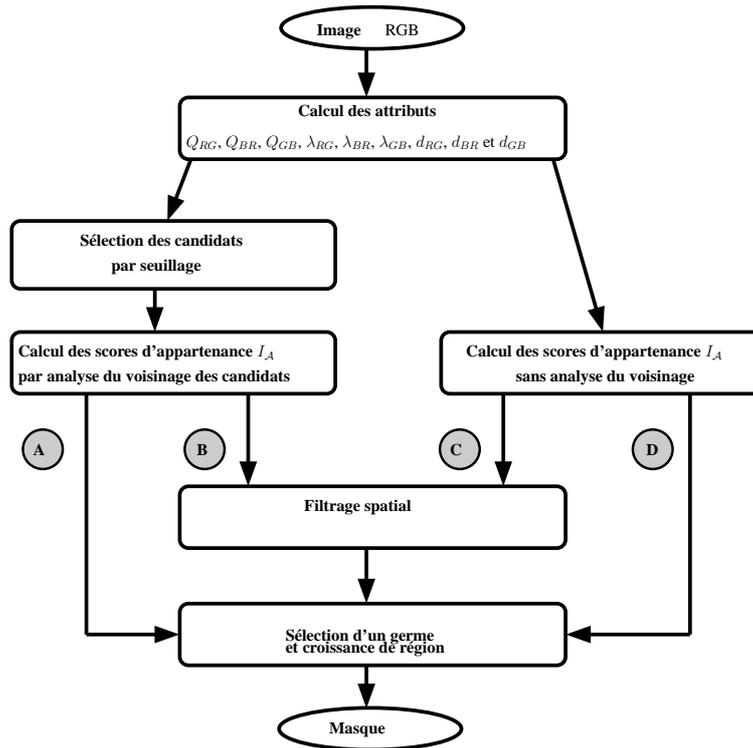


FIG. 4.10 – Structure globale de la méthode de segmentation d'un marquage à l'encre. Quatre variantes sont considérées, selon qu'elles se basent sur la sélection de pixels candidats (méthodes A et B) ou non (méthodes C et D), et qu'elles comprennent un filtrage spatial (méthodes B et C) ou non (méthodes A et D).

4.4 Exemples de segmentation

Tout l'intérêt de l'approche proposée est de détecter une tache d'encre dont la concentration varie au sein du marquage et ce quelle que soit la couleur du fond sur lequel il est imprimé. Prenons l'exemple des images des figures 4.11a 4.11b et 4.11c, où le marquage à l'encre rouge doit être détecté. La première étape consiste au calcul de la couleur du fond C_f . L'analyse d'histogrammes présentée dans la section 4.3.1 permet d'extraire une grande partie du fond afin d'en calculer la couleur moyenne. Les zones de fond extraites avec $\mathcal{S}_h = 5\%$ sont montrées sur les figures 4.11d, 4.11e 4.11f. La valeur du seuil \mathcal{S}_h n'a que peu d'influence sur le calcul de la couleur du fond, dès lors que celui-ci est relativement uniforme. En guise d'exemple, le tableau 4.7 montre les coordonnées obtenues pour $\mathcal{S}_h = 5\%$, 10% et 20% . Les couleurs de fond restent sensiblement similaires.

Nous présentons ici les résultats de quatre types de segmentation, obtenus en exploitant les coefficients de concentration et la distance par rapport au *locus* de l'encre. Ces traitements se différencient entre eux soit par la technique utilisée soit par les attributs pris en compte (attributs marginaux ou globaux, utilisation d'un seul plan dichromatique). La première approche sélectionne les candidats par l'approche présentée dans la section 4.3.2.1 et les trois autres par l'approche de la section 4.3.2.2. Notons que dans toutes ces expérimentations, les paramètres mis en jeu sont identiques.

Fond	C_i	S_h		
		5%	10%	20 %
Vert	R	146.75	145.05	145.13
	G	205.11	201.59	201.80
	B	100.96	99.26	99.37
Jaune	R	202.82	202.53	201.51
	G	163.30	163.02	161.46
	B	50.94	50.82	50.37
Gris	R	181.79	181.70	181.47
	G	186.26	186.19	186.12
	B	179	178.95	178.71

TAB. 4.7 – Couleurs des fonds vert, jaune et gris obtenues avec des seuils S_h de 5%, 10% et 20 %.

Segmentation par multi-seuillage : méthode de type B. Pour chacune des images initiales 4.11a, 4.11b et 4.11c, les trois coefficients de concentration Q_R , Q_G et Q_B ainsi que les trois distances au locus d'encre sont calculés puis seuillés d'après les critères (4.12) et (4.14), en fixant S_s à 50 (la dynamique des images est de 255 niveaux) et les seuils S_{Q_i} à 0.4. Les pixels candidats sont alors sélectionnés à l'aide du critère \mathcal{D}_{Cand_1} donné par (4.15) (ET logique). Ils sont montrés sur les figures 4.12a, 4.12b et 4.12c. Comme nous pouvons le remarquer, cette étape de sélection des pixels candidats n'est pas suffisante pour discriminer totalement le marquage des autres éléments de l'image. Une analyse de voisinage basée sur l'étude d'un critère \mathcal{D}_{app} (voir section 4.3.3.2), permet d'affecter des scores importants aux pixels du marquage, et plus particulièrement pour les pixels de forte concentration. Les paramètres choisis sont $\mu_Q = 0,9$, $\sigma_Q = 0,7$ et $\sigma_D = 50$. Le critère d'appartenance est calculé par combinaison linéaire de \mathcal{D}_Q et \mathcal{D}_d (voir l'équation (4.19) avec $a = c = 0,5$ et $b = 0$) et une taille du voisinage 3×3 . Les images de scores d'appartenance I_A ainsi obtenues sont montrées sur les figures 4.12d, 4.12e et 4.12f. Ensuite, le filtrage spatial expliqué dans la section 4.3.4 est opéré avec $M_x = M_y = 8$. Les images filtrées résultantes I_F sont données sur les figures 4.12g, 4.12h et 4.12i. Le filtrage assure une meilleure connexité des pixels de scores d'appartenance élevés. Finalement, les figures 4.12j, 4.12k et 4.12l montrent le masque englobant le marquage à l'encre. Il est détecté quelle que soit la couleur du fond. Dans toutes les exemples de segmentation décrits ci-dessous, les paramètres mis en jeu sont inchangés.

Utilisation d'un seul plan chromatique (méthodes de types C et D). Pour tous les fonds (vert, jaune, gris) le plan le plus approprié pour calculer les attributs colorimétriques est le plan RG , puisqu'il offre une meilleure disparité entre les couleurs d'encre et de fond. Les critères de distance \mathcal{D}_d , calculés dans le plan RG , sont montrés sur les figures 4.13a, 4.13b et 4.13c, respectivement pour les fonds vert, jaune et gris. Certes, les valeurs de critères sont élevées pour les pixels du marquage mais également pour quelques autres pixels. Les critères de concentration \mathcal{D}_Q , calculés là encore sur les composantes R et G , sont représentés sur les figures 4.13d, 4.13e et 4.13f. À eux seuls ils ne permettraient pas de détecter le marquage. Enfin, les critères \mathcal{D}_1 sont

calculés. Ils sont visibles sur les figures 4.13*g*, 4.13*h* et 4.13*i*. L'analyse de voisinage décrite dans le paragraphe 4.3.3 n'est pas effectuée, de sorte que l'image des scores d'appartenance I_A est directement obtenue par (4.21). Après filtrage, sélection du germe et croissance de région, on aboutit finalement aux masques associés aux marquages à l'encre montrés sur les figures 4.13*j*, 4.13*k* et 4.13*l* (méthode de type C). Les figures 4.13*m*, 4.13*n* et 4.13*o* montrent les résultats obtenus sans filtrage spatial préalable (méthode de type D). Notons que la croissance de région s'effectue dans tous les cas avec $S_{croiss} = 15\%$. Dans le cas où un filtrage est effectué, l'algorithme de détection mis en œuvre a permis de sélectionner un marquage à l'encre quelle que soit la couleur du fond sur lequel il est effectué, malgré quelques imprécisions de détection. Dans le cas où le filtrage n'est pas opéré, les risques de mauvaise détection sont plus importants, comme le montre la figure 4.13*m*, où le marquage n'est pas du tout détecté. En effet, le germe sélectionné correspond dans ce cas à du bruit, que le filtrage spatial permet d'écartier. Ainsi, l'utilisation du filtrage spatial, faute d'améliorer la qualité de la détection (comme nous pouvons le voir en comparant les figures 4.13*l* et 4.13*o*), permet toutefois une meilleure robustesse de la détection.

Utilisation des attributs marginaux Q et d_E (méthodes de type C et D). Ici, seuls sont pris en compte le coefficient de concentration marginal Q et la distance euclidienne d_E par rapport au *locus* du marquage. Là encore, tous les paramètres sont fixés aux mêmes valeurs que précédemment. Les figures 4.14*a*, 4.14*b* et 4.14*c* correspondent au critère \mathcal{D}_d et les figures 4.14*d*, 4.14*e* et 4.14*f* au critère \mathcal{D}_Q calculés respectivement sur les fonds vert (figure 4.11*a*), jaune (figure 4.11*b*) et gris (figure 4.11*c*). Le critère global \mathcal{D}_1 est quant à lui représenté sur les figures 4.14*g*, 4.14*h* et 4.14*i*. L'analyse du voisinage n'étant pas mise en œuvre, on a directement : $I_A = \mathcal{D}_1$. Les résultats obtenus avec filtrage de I_A sont visibles sur les figures 4.14*j*, 4.14*k*, 4.14*l*, tandis que les résultats obtenus sans filtrage spatial préalable sont montrés sur les figures 4.14*m*, 4.14*n*, 4.14*o*. Là encore, les trois masques associés aux marquages sont détectés sur les trois images en utilisant les mêmes paramètres de segmentation. En ce qui concerne les fonds vert et jaune, l'utilisation des attributs marginaux améliore nettement la précision de la segmentation, en comparaison à l'étude du plan couleur RG . En outre, le filtrage spatial n'est pas nécessaire, puisqu'il n'apporte pas d'améliorations sensibles. Dans le cas du fond gris (images 4.14*l* et 4.14*o*), les résultats obtenus sont moins satisfaisants que dans l'expérimentation précédente. Le choix des attributs à utiliser dépend donc fortement du type d'images considéré.

Utilisation des trois composantes couleur (méthodes de type C et D). Pour finir, nous exploitons les attributs calculés sur les trois plans chromatiques RG , BR et GB . Les critères de distance \mathcal{D}_d obtenus respectivement à partir des images 4.11*a*, 4.11*b* et 4.11*c* sont visibles sur les figures 4.15*a*, 4.15*b*, 4.15*c*. Ils sont calculés selon la relation (4.17). Les figures 4.15*d*, 4.15*e* et 4.15*f* correspondent aux critères de concentration \mathcal{D}_Q exprimés par (4.16). Enfin, le critère \mathcal{D}_1 , calculé à partir des critères \mathcal{D}_d et \mathcal{D}_Q , sont représentés sur les figures 4.15*g*, 4.15*h* et 4.15*i*. À ce stade, un filtrage spatial peut être opéré (figures 4.15*j*, 4.15*k* et 4.15*l*) ou non (figures 4.15*m*, 4.15*n* et 4.15*o*). Remarquons que dans ce cas de figure, le filtrage spatial n'apporte pas d'améliorations au niveau de la détection. Dans le cas du fond vert, la segmentation est plus précise sans filtrage (figure 4.15*m*) qu'avec (figure 4.15*j*).

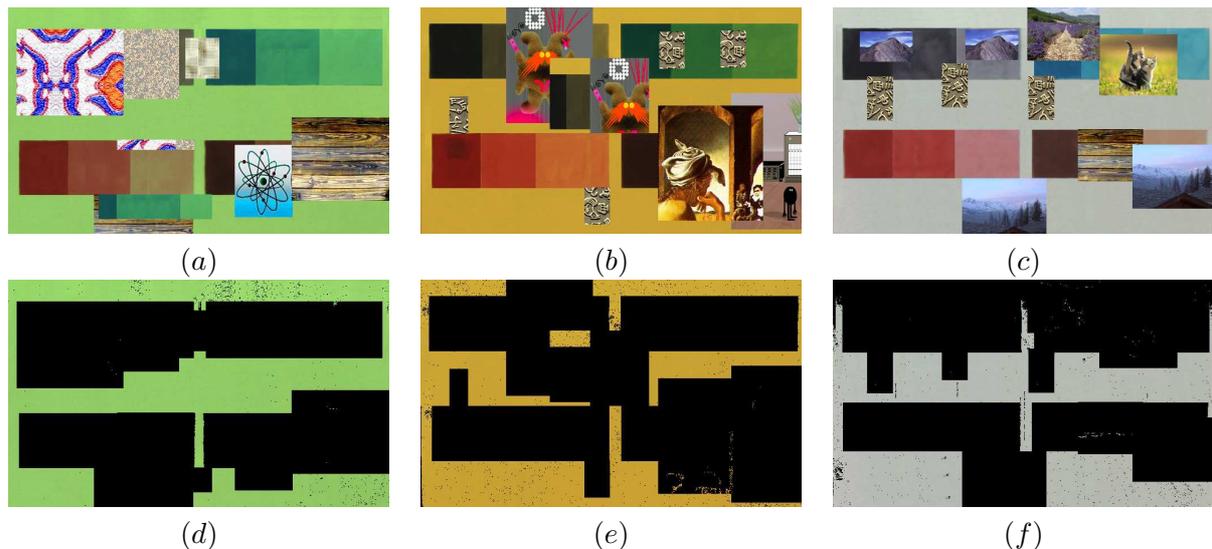


FIG. 4.11 – Détection des fonds pour le calcul de sa couleur.

Discussion. À l'issue de ces quelques exemples de segmentation d'un marquage à l'encre, nous pouvons souligner quelques points importants. Premièrement, le filtrage spatial permet dans tous les cas d'assurer une meilleure robustesse de la détection et une meilleure connexité des pixels constitutifs du marquage. Par contre, il peut nuire à la précision de la segmentation. Ensuite, nous avons vu à travers ces exemples que l'utilisation d'un seul plan couleur (voir l'exemple de la figure 4.13*m* montrant une segmentation à partir du plan RG) ne suffit pas toujours à détecter correctement le marquage. D'autre part, l'utilisation des paramètres globaux aboutit à une moins bonne précision de détection que par l'utilisation des trois plans couleur. Pour nous en convaincre, comparons les résultats des figures 4.12 ou 4.15 (obtenus en utilisant les trois plans couleur) avec les résultats de la figure 4.14 (obtenus avec les composantes globales). Le choix de la méthode de sélection des candidats, par seuillage ou par analyse d'un critère, est assez critique et peut dépendre de l'image à traiter. Comparons à ce propos les figures 4.12*m* 4.12*n* et 4.12*o* (obtenues par seuillage) avec les figures 4.15*m* 4.15*n* 4.15*o* (obtenues par analyse d'un critère). Pour les images de fond vert et jaune, les résultats sont meilleurs après calcul d'un critère que dans le cas d'un multi-seuillage. Or l'inverse est observé pour l'image de fond gris. Le choix de la technique de sélection des candidats dépend donc de l'image à traiter. Toutefois, notons que les paramètres mis en jeu dans les approches sont restés identiques pour toutes les expériences.

4.5 Conclusion

Ce chapitre a abordé le problème de la segmentation d'un marquage à l'encre. Tout d'abord, nous avons montré que la variabilité de la couleur de l'encre imprimée sur un support de couleur se ramène à un problème de transparence. Du fait de ce phénomène, la couleur d'un marquage à l'encre dépend à la fois de la couleur du support sur lequel est imprimée l'encre, et de l'opacité

de celle-ci, qui est décrite par le modèle de Metelli [Metelli, 1970]. L'étude de la transparence de l'encre a permis d'extraire des attributs invariants qui lui sont dédiés, tels que les *coefficients de concentration marginaux* ou le *coefficient de concentration global*. Selon le modèle de convergence de la transparence, ce dernier attribut correspond à une mesure de l'opacité. Ils sont en théorie invariants à la couleur du fond sur lequel est imprimée l'encre. Ensuite, les rapports λ des coefficients de concentration marginaux sont invariants à la quantité d'encre apposée sur le support ainsi qu'aux couleurs de l'encre et du support. La pertinence de ces caractéristiques couleur a été prouvée dans le cadre de la segmentation d'un marquage.

Les quatre algorithmes proposés sont peu dépendants de la couleur de l'encre à segmenter mais il suffit d'effectuer un calibrage au préalable pour la caractériser. De la même manière, ce calibrage permet d'assurer la répétabilité de l'algorithme quelle que soit la couleur de l'illuminant et quelle que soit la sensibilité spectrale des capteurs. Par conséquent, il peut aisément être étendu à diverses applications. Le choix de l'algorithme se justifie d'une part par le type d'application, d'autre part par la tolérance désirée au niveau de la sélection des pixels candidats. Selon l'application, il n'est pas nécessaire d'exploiter tous les attributs décrits, les plus pertinents pouvant être sélectionnés au préalable. Ces algorithmes se veulent génériques et dépendent d'un certain nombre de paramètres qui sont déterminés pour l'instant de manière supervisée par apprentissage sur une base d'images. Une amélioration intéressante de ces techniques serait de rendre plus automatique la détermination de ces paramètres. D'autre part, il pourrait être intéressant d'utiliser d'autres techniques de fusion [Bloch, 2003] pour la sélection des candidats.

Dans le chapitre qui suit, nous exploitons les attributs invariants et la méthode de segmentation décrits dans ce chapitre dans le cadre du contrôle de la traçabilité des produits porcins. Plus précisément, il s'agit de détecter un identifiant imprimé à l'encre sur des pièces de porc, l'objectif étant d'obtenir un masque englobant l'ensemble des pixels du marquage. De manière à justifier l'intérêt de cette approche, nous la comparons à une méthode de segmentation effectuée dans un espace couleur classique.

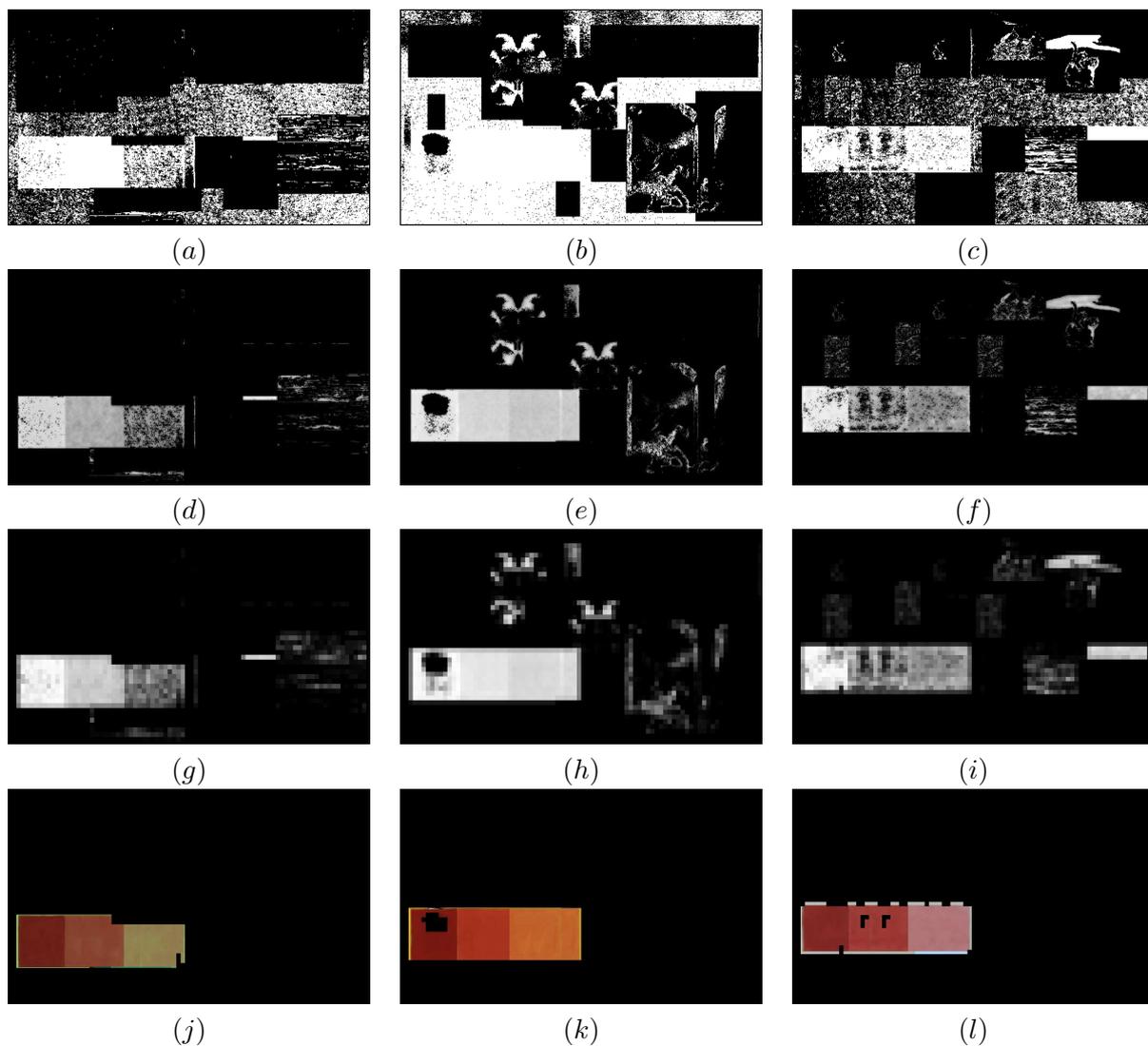


FIG. 4.12 – Exemple de segmentation d'un marquage à l'encre sur trois images différentes, montrant trois fonds différents. Utilisation de multi-seuillage pour la sélection des candidats. (a), (b) et (c) Images des pixels candidats. (d), (e) et (f) Images après analyse de voisinage. (g), (h), (i) Images après filtrage spatial. (j), (k) et (l) Résultats de détection du marquage.

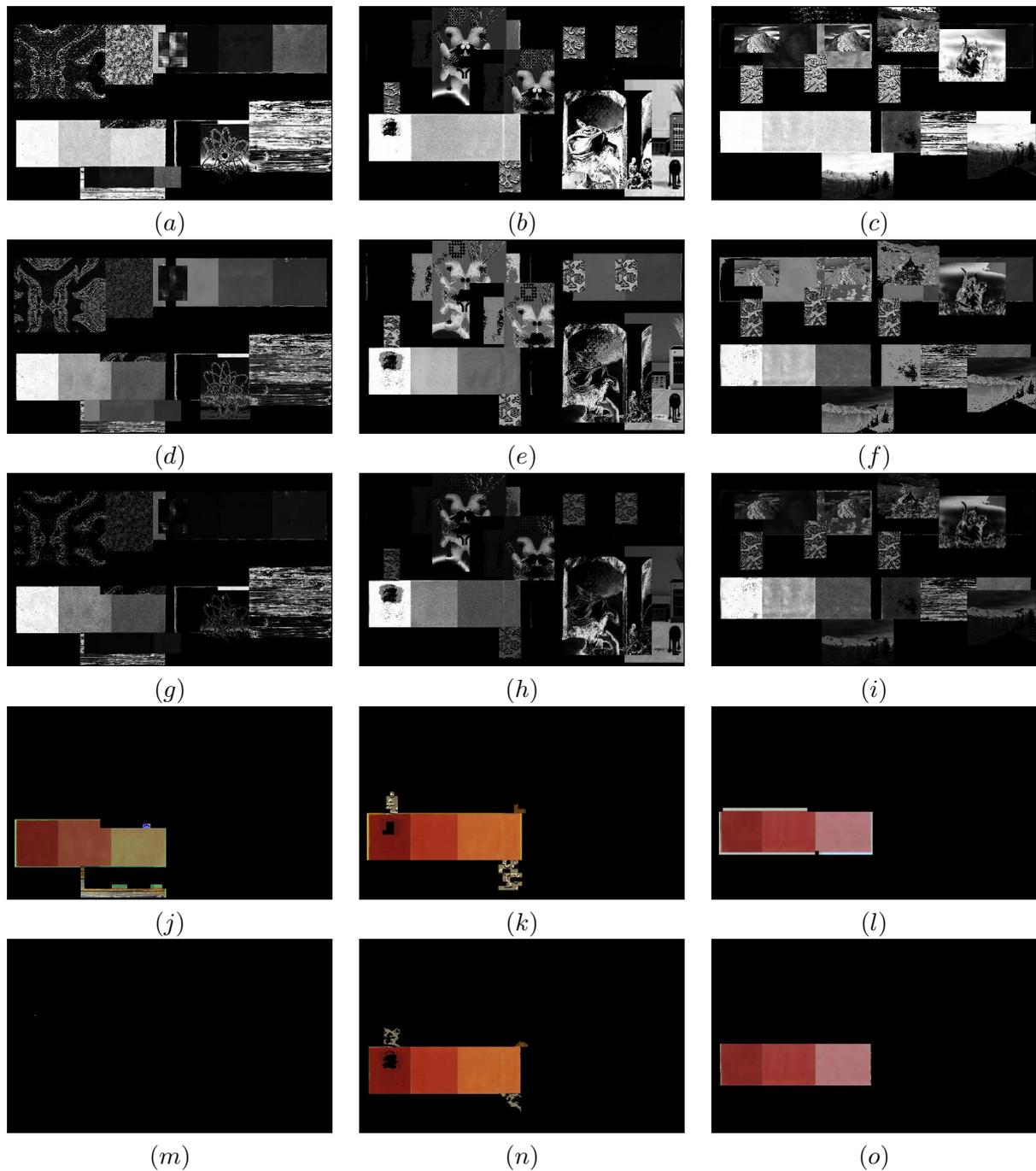


FIG. 4.13 – Exemple de segmentation d'un marquage à l'encre sur trois images différentes, montrant trois fonds différents. Utilisation d'un seul plan RG. (a), (b) et (c) Critères de distance par rapport au locus d'encre \mathcal{D}_d . (d), (e) et (f) Critères de concentration \mathcal{D}_Q . (g), (h), (i) Critères globaux \mathcal{D}_1 . (j), (k) et (l) Résultats de la croissance de région après un filtrage spatial. (m), (n) et (o) Résultats de la croissance de région après un filtrage spatial.

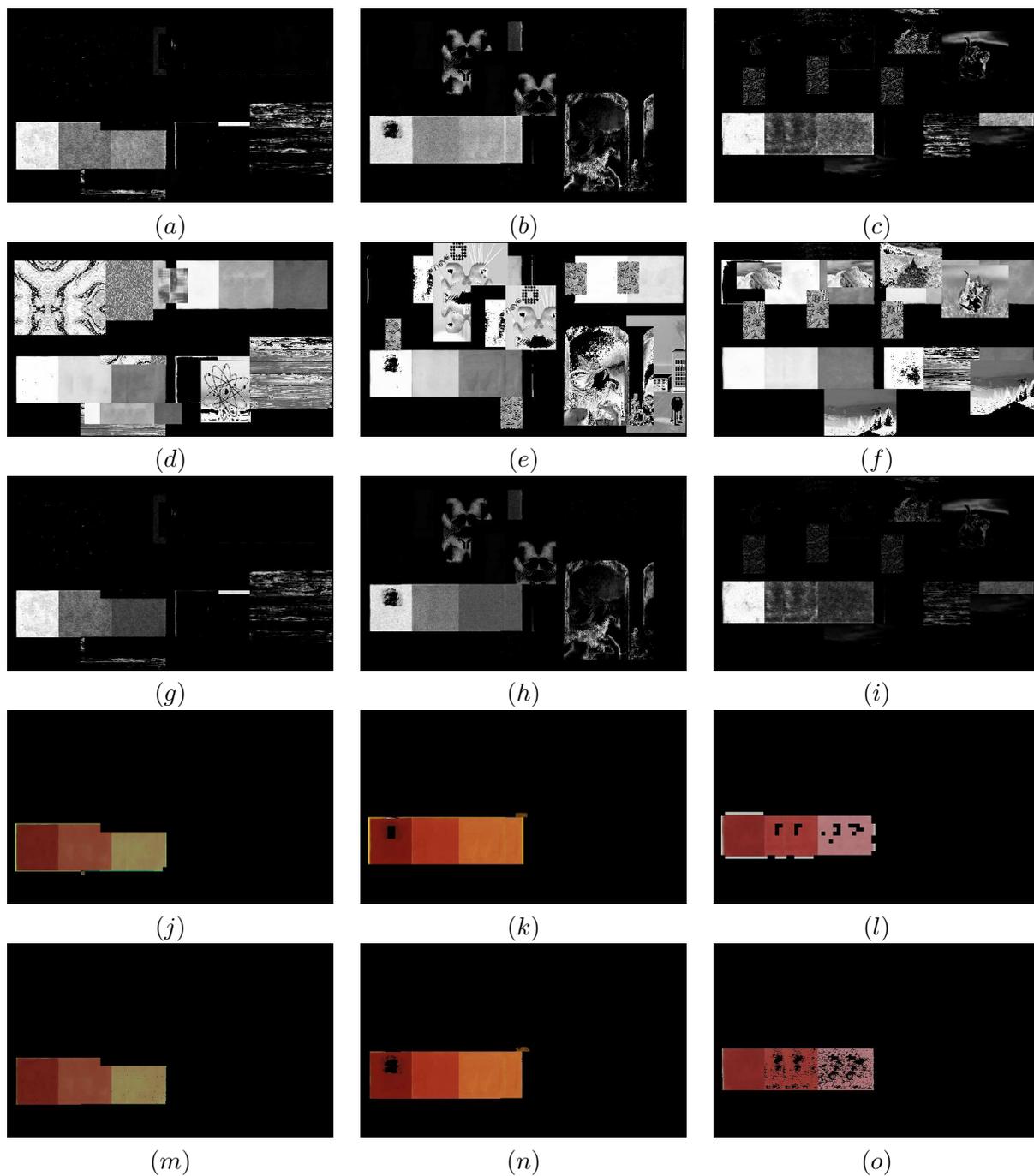


FIG. 4.14 – Exemple de segmentation d'un marquage à l'encre sur trois images différentes, montrant trois fonds différents. Utilisation du coefficient de concentration marginal Q et de la distance euclidienne d_E par rapport au locus du marquage. (a), (b) et (c) Critères de distance par rapport au locus d'encre \mathcal{D}_d . (d), (e) et (f) Critères de concentration \mathcal{D}_Q . (g), (h), (i) Critères globaux \mathcal{D}_1 . (j), (k) et (l) Résultats de la croissance de région après un filtrage spatial. (m), (n) et (o) Résultats de la croissance de région après un filtrage spatial.

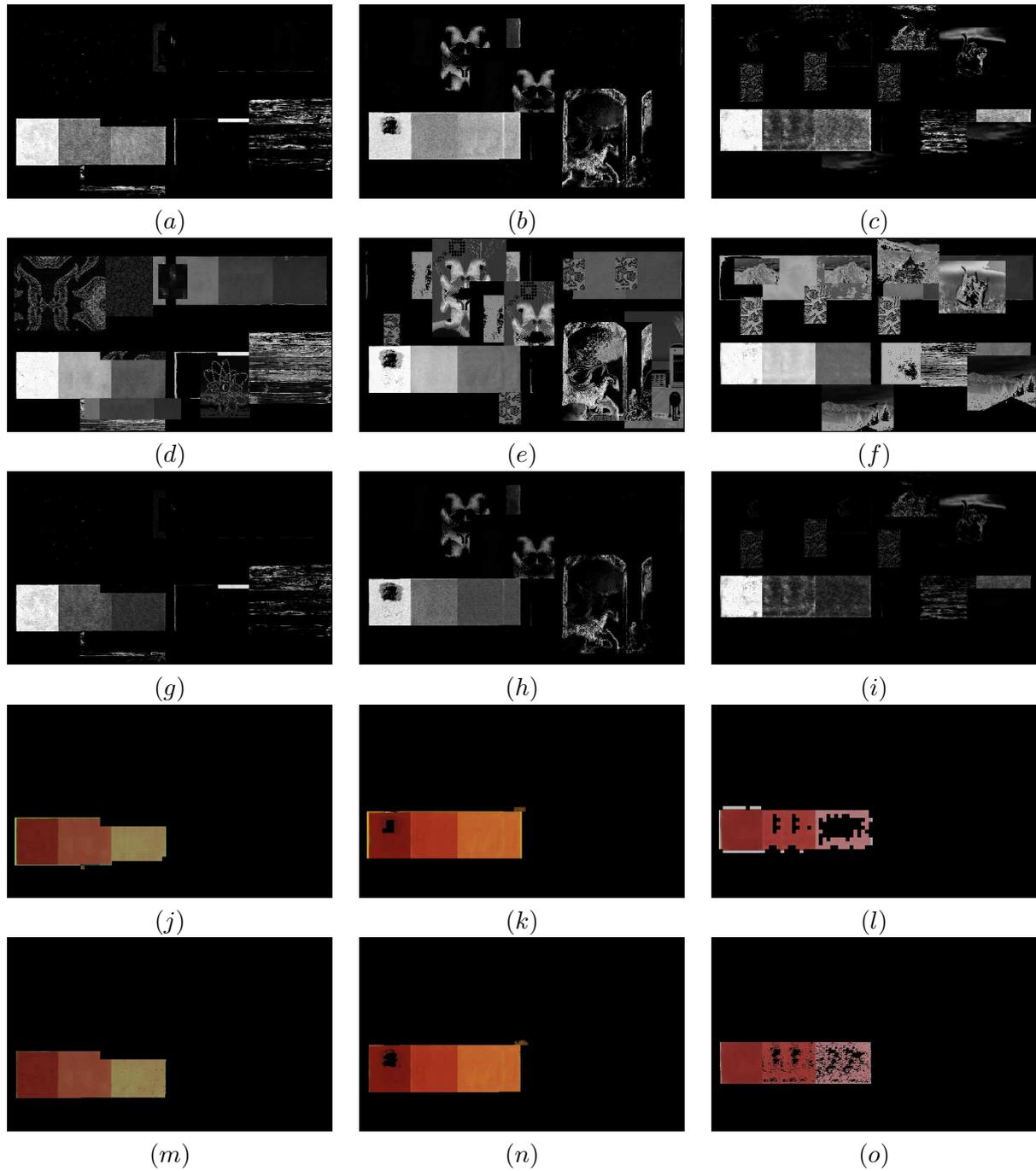


FIG. 4.15 – Exemple de segmentation d'un marquage à l'encre sur trois images différentes, montrant trois fonds différents. Utilisation des trois composantes couleur. (a), (b) et (c) Critères de distance par rapport au locus d'encre \mathcal{D}_d . (d), (e) et (f) Critères de concentration \mathcal{D}_Q . (g), (h), (i) Critères \mathcal{D}_{CandC} . (j), (k) et (l) Résultats de la croissance de région après un filtrage spatial. (m), (n) et (o) Résultats de la croissance de région sans filtrage spatial.

APPLICATION AU CONTRÔLE DE LA TRAÇABILITÉ PORCINE

Ce chapitre illustre l'apport des attributs invariants et de la méthode de segmentation proposés dans le chapitre 4, en les appliquant dans un contexte de contrôle de la traçabilité de produits porcins. Il s'agit de détecter un identifiant imprimé à l'encre sur les jambons, en vue de sa reconnaissance. Afin de prouver l'intérêt de cette approche, nous la comparons à une technique de segmentation effectuée dans un espace couleur classique [Gouiffès et al., 2004c], où la détection robuste du marquage passe par une analyse de texture.



Du fait de récents problèmes d'hygiène alimentaire dans la chaîne de consommation de produits carnés, la traçabilité est aujourd'hui une priorité. Il s'agit en effet du seul moyen permettant de rassurer un consommateur de plus en plus méfiant.

Notre travail s'inscrit dans cette problématique, puisqu'il s'agit de garantir le contrôle de la traçabilité de produits porcins à l'intérieur et à la sortie de l'abattoir. La filière porcine a mené des études, en terme d'efficacité et d'économie, sur les moyens de garantir une identification fiable des animaux, des carcasses puis des différentes pièces de viande. Le tatouage et le marquage à l'encre sont les solutions retenues. Tout lieu d'élevage porcine est identifié par un *numéro de cheptel* reçu de l'Établissement Départemental de l'Élevage. Ensuite, chaque porcelet destiné à l'abattoir reçoit un *indicatif de marquage*, qui lui est tatoué à la naissance. Après abattage, deux marquages supplémentaires sont imprimés par jet d'encre sur la carcasse : le *cachet vétérinaire* et le *numéro d'abattage*, normalisé par l'AFNOR depuis 1999. Il se compose de sept chiffres : les trois premiers correspondent au quantième du jour d'abattage, tandis que les quatre suivants sont associés au numéro d'abattage du jour considéré. De cette manière, un numéro unique est attribué à une carcasse (chaque année). Cet identifiant permet ensuite d'accéder à l'information de provenance de l'animal. Notons que les couleurs d'encre utilisées pour l'impression peuvent varier d'un abattoir à l'autre.

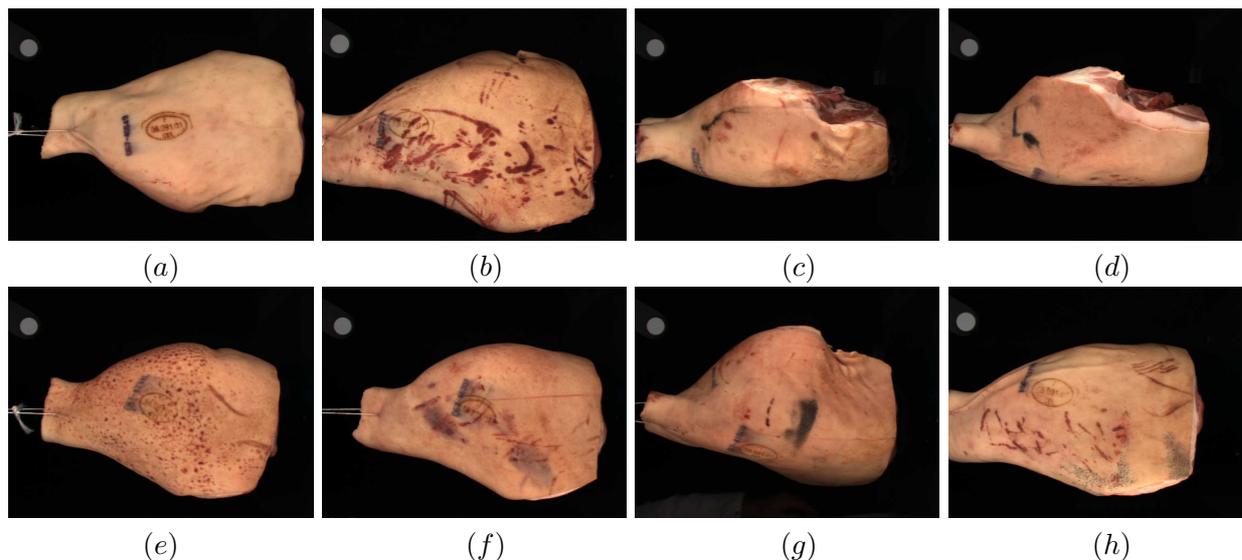


FIG. 5.1 – Exemples d’images à traiter. (a) Jambon normal. (b) Brûlures. (c) et (d) Présence de chair et occultation partielle de l’identifiant. (e) érythèmes. (f) Hématomes. (g) Taches de graisse. (h) Poils.

Généralement, le numéro d’abattage est ensuite retranscrit par impression sur les huit pièces de découpe principale : jambons, épaules, poitrines, longes. Ces pièces sont entreposées dans l’abattoir dans l’attente de leur expédition. Avant leur sortie de l’abattoir, l’identifiant doit être lu afin que la traçabilité soit assurée à chaque stade de la transformation du produit. Cette tâche est à ce jour effectuée par un opérateur. Actuellement, elle n’est que partiellement assurée. Le Cemagref, en partenariat avec l’OFIVAL¹, à travaillé sur la mise en œuvre d’un système de vision numérique, permettant la détection puis la lecture automatique de cet identifiant. Le travail relaté dans ce chapitre concerne plus particulièrement la détection du marquage. Celui-ci doit ensuite être utilisé comme primitive dans une tâche de positionnement de la caméra de manière à en améliorer la lisibilité [Collewet et Chaumette, 2002, Alhaj, 2004]. Enfin, un algorithme de reconnaissance de caractères doit pouvoir assurer sa lecture.

La figure 5.1 montre quelques exemples de jambons, où le numéro d’abattage de l’animal est formé de caractères numériques de couleur noire bleutée. La principale difficulté est de différencier l’identifiant des autres marques présentes sur la couenne. Dans la plupart des cas, le jambon apparaît sans défaut (figure 5.1a). Néanmoins, les brûlures (figure 5.1b), la présence de chair ou l’occultation partielle de l’identifiant (figure 5.1c), les érythèmes (figure 5.1d), les hématomes (figure 5.1e), les taches de graisses (figure 5.1f), ou encore les poils (figure 5.1h) sont autant d’éléments susceptibles de compliquer la détection. Les bavures de l’encre, les déformations dues à la fois à la courbure du jambon et au mode d’impression, mais aussi la présence d’autres marquages tels que le cachet vétérinaire (de forme ovale sur les images de la figure 5.1), interdisent la seule utilisation d’une méthode de reconnaissance de caractères ou encore l’analyse de critères de forme. Ajoutons à cela l’hétérogénéité de la couleur de la couenne. Pour illustrer ce propos, le tableau 5.1, présente les moyennes RGB des couleurs de couennes calculées sur 200 jambons environ ainsi que leur valeurs maximale et minimale.

¹Office national Interprofessionnel des Viandes, de l’élevage et de l’aviculture.

TAB. 5.1 – Mesures des couleurs de couenne

	R	G	B
min	160	87	60
max	210	161	111
moy	186	124	86

L'approche de segmentation proposée dans le chapitre 4, en se basant sur des attributs invariants vis-à-vis de la concentration en encre et des attributs invariants vis-à-vis de la couleur du fond, semble tout à fait appropriée à cette problématique. Après avoir présenté le système de vision utilisé (section 5.1), la section 5.2 explique la mise en œuvre des approches de segmentation utilisant les attributs invariants dédiés à l'encre. Afin d'évaluer la pertinence de l'utilisation d'invariants, nous comparons ces techniques avec une méthode de segmentation mise en œuvre dans un espace couleur classique [Gouiffès *et al.*, 2004c]. Enfin, les résultats de segmentation sont étudiés dans la section 5.4.

5.1 Système de vision et base d'images

La reflectance du jambon, dont la mesure est expliquée dans l'annexe A, a montré l'aspect très spéculaire de cet objet. Puisque l'apparition de spécularités peut nuire aux tâches de traitement d'images, il est nécessaire soit d'éliminer ces spécularités, soit de considérer un système de vision permettant de réduire ces artefacts. Nous avons opté pour la seconde solution, en utilisant un système d'éclairage diffus et uniforme développé au Cemagref dans le cadre d'une application de détection de défauts sur des produits porcins [Marty-Mahé *et al.*, 2003]. Ce système est représenté sur la figure 5.2. Certes son encombrement peut rendre difficile son utilisation dans un cadre industriel, ou encore dans le cadre de l'asservissement visuel où la caméra doit se déplacer en face de l'identifiant. L'étude de la reflectance du jambon (annexe A) nous a permis d'envisager une configuration d'éclairage permettant de réduire l'influence des spécularités. Il pourrait être obtenu en considérant plusieurs sources d'éclairage direct portées par la caméra, et dont les directions d'éclairage forment un angle inférieur à 30 avec la direction d'observation.

Une caméra triCCD Sony DXC 990 P et un logiciel d'acquisition d'images fournissent des images *RGB* codées sur huit bits par bande. Nous disposons ainsi d'une base de plus de 200 images de jambons, montrant pour la plupart des défauts. Notons que la caméra est toujours positionnée approximativement à une même distance du jambon.



FIG. 5.2 – Système d’acquisition utilisé.

5.2 Approches de segmentation par utilisation des invariants dédiés à l’encre

La variabilité de la quantité d’encre utilisée ainsi que la grande variété de couleurs de couennes rencontrées (voir le tableau 5.1) constituent des difficultés qui peuvent être largement simplifiées par l’utilisation de l’algorithme de segmentation présenté dans le chapitre 4. Les coordonnées couleur de l’encre utilisée sont $R = 20$, $G = 10$ et $B = 30$. Quant à la couleur du fond, elle est calculée par la méthode d’analyse d’histogramme présentée dans la section 4.3.1.

Notons de prime abord que, comme l’atteste la figure 5.3 illustrant la distribution des pixels de marquage dans l’espace RGB , les plans chromatiques RG (image 5.3b), BR (image 5.3c) et GB (image 5.3c) montrent tous une bonne disparité entre la couleur de l’encre et la couleur de la couenne. Ils sont donc peu susceptibles d’être bruités. Ainsi, l’ensemble de ces attributs est exploité. La figure 5.5 montre quelques uns des attributs invariants calculés sur l’image 5.4. L’image des distances euclidiennes d_E , du coefficient de concentration marginal Q , et le rapport des coefficients de concentration λ^{RG} sont respectivement visibles sur les figures 5.5a, 5.5b et 5.5c. Remarquons notamment que les pixels du marquage sont proches colorimétriquement de la droite $[C_e, C_f]$ (faible intensité dans l’image 5.5a) et qu’ils détiennent un fort coefficient de concentration (forte intensité dans l’image 5.5b).

Plusieurs types d’algorithmes de segmentation d’un marquage à l’encre (A, B, C et D sur la figure 4.10) ont été proposés dans le chapitre précédent. Étant donnée la nature de l’identifiant à détecter, constitué de pixels qui ne sont pas tous connexes, le filtrage spatial est requis. Cela nous amène donc à privilégier les deux variantes B et C dans le cadre de la traçabilité des produits carnés. Néanmoins, il convient tout d’abord d’éliminer les pixels associés aux ombres et aux taches de graisse, qui sont susceptibles de nuire à la segmentation de l’identifiant.

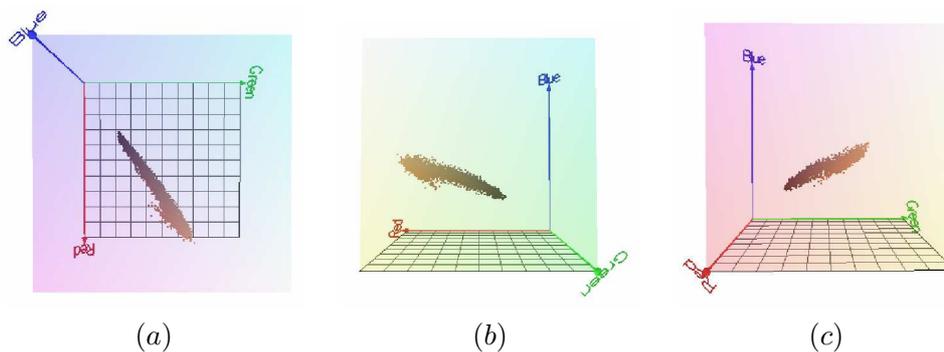


FIG. 5.3 – Exemple de distribution des pixels du marquage noir-bleuté sur la couenne, dans l'espace RGB.

5.2.1 Extraction des ombres à l'aide des rapports de coefficients de concentration

Les ombres, dues à la courbure et aux plis du jambon ainsi que les taches de graisse présentes sur certaines pièces de porc ont parfois des caractéristiques couleur très proches du marquage à l'encre, et sont donc susceptibles de nuire à sa détection. D'après le modèle dichromatique [Shafer, 1985], abordé dans la section 1.2.2.6, les couleurs associées à un matériau lambertien de couleur de réflexion diffuse *body* C_b (ou couleur de *body*) sont toutes situées sur le segment $[O C_b]$ (voir l'annexe B). Si l'on considère que, dans le cas des images de jambon, l'éclairage diffus permet d'annuler l'influence de la réflexion spéculaire alors $C_b = C_f$. Par conséquent, les couleurs associées aux ombres sont situées sur le segment $[O C_f]$. De même, les taches de graisse de couleur noire se situent sur le segment $[O C_f]$ et les attributs de transparence de l'encre peuvent être transposés au cas de la transparence des taches de graisse. Notamment, il est possible de leur associer des rapports de coefficients de concentration, que nous appelons λ_g^{RG} , λ_g^{BR} et λ_g^{GB} . Les ambiguïtés possibles entre l'identifiant d'un côté et les ombrages et taches de graisse d'un autre côté, peuvent être levées en partie par comparaison des coefficients de graisse λ_g^{RG} , λ_g^{BR} , λ_g^{GB} et les coefficients associés à l'encre λ_e^{RG} , λ_e^{BR} , λ_e^{GB} . Ainsi, les pixels p dont les coefficients d'encre sont plus proches de 1 que ne le sont les coefficients de graisse appartiennent plus probablement au segment $[C_e C_f]$ qu'au segment $[O C_f]$. La première étape de la segmentation consiste ainsi à détecter les pixels satisfaisant le critère \mathcal{C}_{ombres} suivant :

$$\mathcal{C}_{ombres} = \begin{cases} 1 & \text{si } |\lambda_e^{RG} - 1| < |\lambda_g^{RG} - 1|, |\lambda_e^{BR} - 1| < |\lambda_g^{BR} - 1| \text{ et } |\lambda_e^{GB} - 1| < |\lambda_g^{GB} - 1| \\ 0 & \text{sinon} \end{cases} \quad (5.1)$$

Sur la figure 5.5c est représentée la classe des pixels obtenue après analyse du critère \mathcal{C}_{ombres} à partir de l'image 5.4. À ce stade, on peut soit opérer une sélection des pixels candidats à l'identifiant (méthode de type de type B), soit directement calculer un score d'appartenance (méthode de type D).

5.2.2 Segmentation avec analyse de voisinage

La méthode abordée ici a été développée dans la section 4.3. Elle correspond à la méthode B présentée sur la figure 4.10. Comme la plupart des approches de segmentation, elle requiert la détermination de paramètres qui peuvent dépendre du type d'image. Ces valeurs ont été définies par apprentissage supervisé sur une vingtaine d'images représentatives de la base.

Sélection des pixels candidats. La sélection des pixels candidats est effectuée à l'aide de la technique de multi-seuillage, comme l'explique la section 4.3.2.1. Elle nécessite de fixer les valeurs \mathcal{S}_Q , \mathcal{S}_λ , \mathcal{S}_d pour le calcul des critères donnés par (4.12), (4.13) et (4.14). Du fait de la relative opacité de l'encre, le seuil \mathcal{S}_Q est fixé à 0.5. \mathcal{S}_d dépend de la disparité des couleurs autour de la droite $[C_e, C_f]$. Nous avons considéré une valeur de 20. En ce qui concerne \mathcal{S}_λ , un écart de 15 % autour de l'unité est toléré. Étant donné le faible écart des couleurs par rapport à $[C_e, C_f]$, ces valeurs sont susceptibles d'être satisfaisantes quelle que soit l'image utilisée. Une fois fixés, ces différents seuils ne sont plus modifiés, quelle que soit l'image de jambon traitée. Sur la figure 5.6, on peut suivre les différentes étapes du traitement de segmentation concernant l'image présentée sur la figure 5.4. Les pixels candidats sont représentés par des pixels blancs sur l'image 5.6a. Remarquons qu'un certain nombre de pixels de l'image 5.5d a ainsi été éliminé. Les *pixels candidats* restants sont principalement situés sur l'identifiant et sur les bords du jambon.

Élimination des faux candidats. Les mauvais candidats sont éliminés par analyse de leur voisinage, comme le décrit la section 4.3.3.2, en calculant une image des scores d'appartenance au marquage I_A dépendant des valeurs des voisins de chaque pixel. Le critère d'appartenance \mathcal{D}_{app} mis en jeu est donné par la relation (4.19), c'est-à-dire par combinaison linéaire des trois critères \mathcal{D}_Q , \mathcal{D}_λ et \mathcal{D}_d exprimés respectivement par (4.16), (4.17) et (4.18). Les valeurs prises en compte dans cette combinaison sont $a = c = 0.25$ et $b = 5$. Quant aux critères \mathcal{D}_Q , \mathcal{D}_λ et \mathcal{D}_d , les paramètres choisis sont $\mu_Q = 1$, $\sigma_Q = 0.4$, $\sigma_L = 0.15$, $\sigma_D = 15$. Enfin, le voisinage de l'analyse est de taille 5×5 . L'image des scores d'appartenance I_A , représentée sur la figure 5.6b, montre de fortes valeurs au niveau du marquage mais également en quelques endroits du bord du jambon.

Sélection d'une composante connexe associée au marquage. De manière à obtenir un masque associé au marquage, l'image I_A est partitionnée en fenêtres W_{uv} , comme cela a été décrit dans la section 4.3.4. Étant donnée la hauteur moyenne des caractères dans l'image (environ 15 à 20 pixels), une fenêtre W_{uv} , de taille 24×24 pixels est suffisante dans tous les cas. L'image $I_{\mathcal{F}}$ obtenue à l'issue de ce traitement est montrée sur la figure 5.6c. La détection de la zone $\mathcal{M}_{W_{MAX}}$ détenant la valeur de score d'appartenance la plus élevée est sélectionnée comme germe. Enfin, la croissance de région est effectuée avec $\mathcal{S}_{croiss} = 40\%$. L'image 5.6d représente l'emplacement du marquage. La zone détectée permet d'englober la totalité des caractères associés à l'identifiant.

Afin d'évaluer la nécessité ou non de l'analyse de voisinage, une seconde approche de segmentation n'incluant pas cette étape est proposée.

5.2.3 Segmentation sans analyse de voisinage

La méthode proposée ici correspond à la technique D sur la figure 4.10, développée dans la section 4.3. Dans un premier temps, un critère \mathcal{D}_1 (voir la section 4.3.2.2 page 84) permet de quantifier l'appartenance de chaque pixel au marquage et d'obtenir directement l'image des scores d'appartenance I_A (voir la relation (4.21) page 85). L'extraction de la composante connexe s'effectue alors à partir de cette dernière image.

Calcul du critère d'appartenance. Le critère d'appartenance au marquage, appelé \mathcal{D}_1 est obtenu par le produit des trois critères, \mathcal{D}_Q , \mathcal{D}_d et \mathcal{D}_λ , calculés en prenant $\sigma_d = 20$, $\sigma_Q = 0.5$, $\mu_Q = 1$ et $\sigma_\lambda = 0.15$. La figure 5.7a montre l'image I_A obtenue.

Sélection de la région associée au marquage. L'extraction de la composante connexe associée à l'identifiant s'effectue de la même manière que pour l'approche précédente (voir 5.2.2). L'image I_A est partitionnée, comme le montre la figure 5.7b. La valeur maximale de cette nouvelle image sert alors de germe pour effectuer la croissance de région. Par contre, du fait de la moins bonne séparabilité entre le marquage et les pixels associés à la couenne, il faut sélectionner un seuil de croissance de région plus strict $\mathcal{S}_{croiss} = 60\%$. Finalement, le masque associé à la région recherchée est montré sur la figure 5.7c.

La prochaine section décrit une méthode de segmentation développée à partir d'un système couleur classique.

5.3 Segmentation dans un espace couleur classique

Comme l'a déjà souligné le chapitre 2, la détermination du système couleur le plus adapté à la segmentation est une tâche difficile. À l'aide du logiciel *Colorspace*², qui permet la visualisation des images dans différents espaces de représentation de la couleur, nous avons comparé un grand nombre d'espaces colorimétriques usuels, vis à vis de la discrimination du marquage par rapport aux défauts et autres marques présents sur la couenne. Nous nous sommes alors orientés vers l'utilisation d'un système *HSV*. Nous effectuons une analyse factorielle discriminante sur les composantes H' , S' et V' , qui représentent respectivement les composantes H , S et V normalisées de manière à s'exprimer sur la même dynamique. L'analyse factorielle discriminante permet de déterminer les coefficients α , β et γ de la combinaison linéaire $\theta = \alpha H' + \beta S' + \gamma V'$ autorisant la plus grande discrimination de l'identifiant vis-à-vis des autres classes présentes dans l'image. Dans le cadre de notre application, $\alpha = \beta = 1$ et $\gamma = 0$. La combinaison linéaire de H et S permet *a priori* de discriminer au mieux le marquage des autres régions présentes dans l'image. D'après l'analyse des espaces couleurs du chapitre 2, la teinte est bruitée pour des saturations proches de 0, et il peut être intéressant dans certaines applications d'utiliser une fonction de pertinence de la teinte, comme pour le gradient de Carron (voir la section 3.1.2.1). On peut également ne pas considérer les teintes des couleurs pour lesquelles la saturation est

²de Philippe Colantoni du LIGIV St Etienne, disponible sur le site <http://www.couleur.org/>



FIG. 5.4 – Image originale.

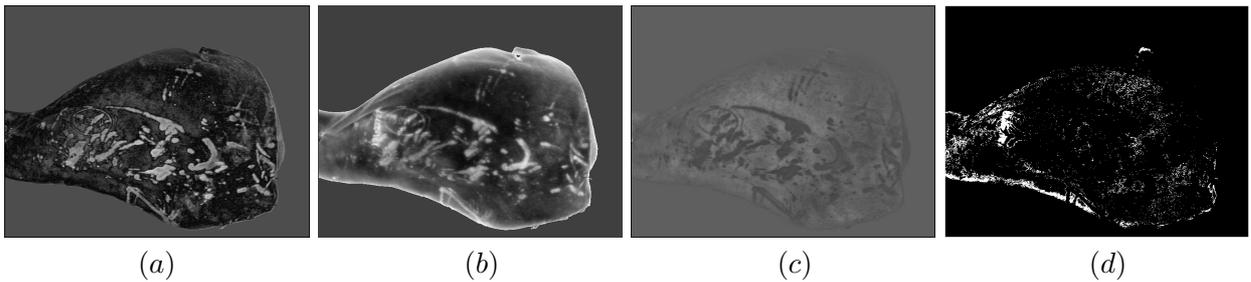


FIG. 5.5 – Calcul des attributs invariants. (a) Distance euclidienne par rapport au locus $[C_e C_f]$. (b) Coefficient de concentration QR . (c) Rapport des coefficients de concentration λ^{RG} . (d) Image des pixels répondant au critère C_{ombres} après extraction des pixels d'ombres.

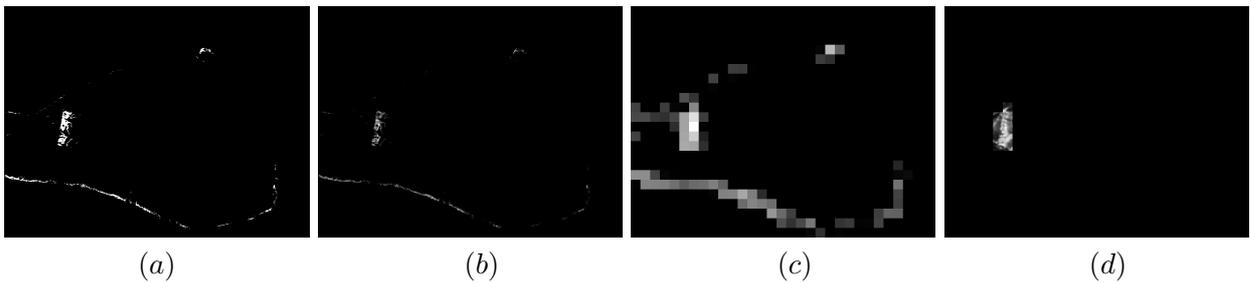


FIG. 5.6 – Illustration de la méthode de segmentation avec analyse de voisinage. (a) Pixels candidats. (b) Image des scores d'appartenance I_A après analyse de voisinage. (c) Image des moyennes M_W après filtrage spatial. (d) : Résultat de la détection.

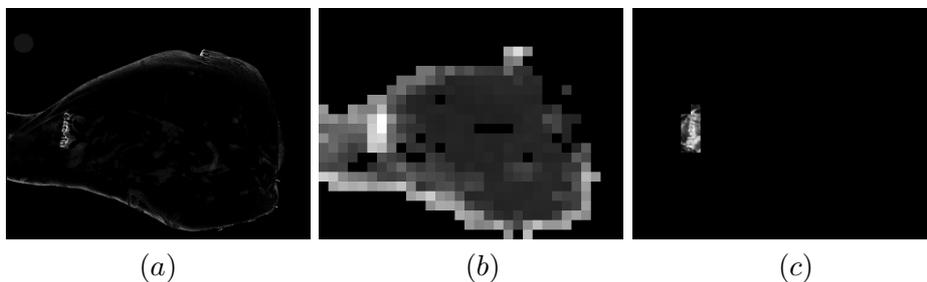


FIG. 5.7 – Illustration de la méthode de segmentation sans analyse de voisinage. (a) Critère d'appartenance. (b) Image I_F obtenue après filtrage spatial. (c) Résultat de la détection.

inférieure à un seuil (voir par exemple [Ohba et Ikeuchi, 2000])³. Dans le cadre de notre application, l'utilisation de cette pondération n'est pas indispensable.

L'algorithme de segmentation proposé débute par la classification spatio-colorimétrique des pixels basée sur l'analyse des degrés de connexité [Fontaine *et al.*, 2000], puis sélectionne le marquage en fonction d'un critère basé sur la texture.

5.3.1 Classification spatio-colorimétrique

Dans la section 3.2.3, nous avons décrit une méthode de classification spatio-colorimétrique proposée dans [Fontaine *et al.*, 2000, Fontaine, 2001]. Elle se base sur le calcul d'une *Pyramide des Degrés de Connexité Couleur*. Cette technique repose sur l'hypothèse selon laquelle chaque segment de l'image est constitué de pixels qui appartiennent à un intervalle couleur de forte connexité.

Soient les coordonnées colorimétriques C_i , pour $i = R, G, B$. L'auteur calcule les *degrés de connexité* (ce terme est explicité plus loin) relatifs à des ensembles de couleurs bichromatiques $[C_R, C_R + l; C_G, C_G + l]$, $[C_R, C_R + l; C_B, C_B + l]$ et $[C_B, C_B + l; C_G, C_G + l]$, avec l la taille de l'intervalle de couleur de telle sorte que $0 < l < L$, où L est la taille maximale de l'intervalle. On se référera à la figure 3.8 pour visualiser cette structure de données. Chaque cellule de la pyramide, de coordonnées (C_i, C_j, l) , contient le degré de connexité de l'intervalle de couleur $[C_i, C_i + l; C_j, C_j + l]$, pour $i \neq j$.

De cette manière, la base de la pyramide contient les degrés de connexité de chaque couleur susceptible d'être présente dans l'image, tandis que le sommet contient le degré de connexité associé à l'intervalle de toutes les couleurs de l'image. Cette classification spatio-colorimétrique, non-supervisée, ne requiert aucune connaissance du nombre de classes présentes dans l'image. Ce point s'avère tout à fait pertinent dans notre cadre d'étude, où le nombre de classes de défauts et marques présentes sur la couenne varie d'un jambon à l'autre. Toutefois, le coût en temps de calcul exorbitant que représente la construction des trois pyramides nous amène à considérer une version simplifiée de cette structure de données. M. Fontaine propose une version unidimensionnelle pour la segmentation d'images en niveaux de gris [Fontaine *et al.*, 2000]. Nous nous inspirons de cette structure pour analyser les degrés de connexité du seul plan couleur θ , explicitée précédemment. Nous appelons *triangle de connexité* (TdC) cette structure de données.

En pratique, les traitements sont restreints aux cellules associées aux intervalles de taille impaire. De cette manière, la cellule de TdC de coordonnées (θ_a, l) contient le degré de connexité de l'intervalle de couleurs $[\theta_a - l, \theta_a + l]$, avec $0 \leq l \leq L$. Ainsi, la ligne d'indice l du TdC contient les degrés de connexité d'intervalles de taille $T_l = 2 \times l + 1$. Étant donné le nombre maximum de couleurs (NbC) de composante θ potentiellement présentes dans l'image (256 dans notre cas) et la taille d'intervalle maximale $T_L = 2 \times L + 1$, la taille du triangle de connexité est donnée par $L \times NbC$. On comprendra facilement qu'il s'avère inutile d'analy-

³On parle alors de teinte achromatique. La limite entre teintes achromatiques et teintes chromatiques est définie par le diagramme de Kelly [Agoston, 1987]. Il y est montré que les couleurs de saturation inférieure à 10 % sont achromatiques.

ser une telle taille d'intervalle de pixels, puisque cela revient à supposer qu'il n'existe qu'une classe de pixels dans l'image. De la même manière, il paraît inutile d'analyser un intervalle de trop grande taille comparativement aux intervalles des classes présentes dans l'image. Nous proposons donc de limiter le calcul du TdC aux cellules associées à des tailles T_l d'intervalles inférieures à un seuil S_l , dont l'ajustement dépend exclusivement du contenu d'image à analyser. Dans le cadre de notre application, S_l est fixé à 40, sachant que la composante θ est codée sur 256 valeurs.

Avant d'explicitier le formalisme de cette structure, revenons sur quelques définitions, en particulier sur la définition de la probabilité de premier ordre d'un ensemble de pixel et la probabilité de cooccurrence couleur. Nous aborderons ensuite la définition du degré de connexité utilisé.

Probabilité de premier ordre d'un ensemble de couleurs. La probabilité de premier ordre $P_1([\theta_a - l, \theta_a + l])$ d'un ensemble de couleurs $[\theta_a - l, \theta_a + l]$ correspond à la somme des probabilités qu'un pixel de couleur θ_i appartienne à l'intervalle de couleurs $[\theta_a - l, \theta_a + l]$. Soit $P_1(\theta_i)$ la probabilité qu'un pixel soit de couleur θ_i . La probabilité de premier ordre de l'ensemble de couleurs $[\theta_a - l, \theta_a + l]$ est alors donnée par :

$$P_1([\theta_a - l, \theta_a + l]) = \sum_{\theta_i=\theta_a-l}^{\theta_i=\theta_a+l} P_1(\theta_i) \quad (5.2)$$

Probabilité du second ordre d'un ensemble de couleurs. La probabilité de cooccurrence couleur $P_{cc}(\theta_a, \theta_b)$ d'un couple de couleurs (θ_a, θ_b) représente la probabilité pour qu'un pixel appartenant au voisinage $V(\theta_a)$ d'un pixel de couleur θ_a au sens du 8-voisinage, soit de couleur θ_b :

$$P_{cc}(\theta_a, \theta_b) = \frac{1}{8} \sum_{\theta_b \in V(\theta_a)} P_{oc}(\theta_a, \theta_b) \quad (5.3)$$

où $P_{oc}(\theta_a, \theta_b)$ représente la probabilité pour que θ_a et θ_b soient les couleurs de deux pixels voisins. La probabilité du second ordre $P_2([\theta_a - l, \theta_a + l])$ d'un intervalle de couleurs $[\theta_a - l, \theta_a + l]$ est alors donnée par la somme des probabilités de cooccurrence couleur de tout couple de couleurs (θ_a, θ_b) appartenant à l'intervalle de couleurs $[\theta_a - l, \theta_a + l]$:

$$P_2([\theta_a - l, \theta_a + l]) = \sum_{\theta_i=\theta_a-l}^{\theta_i=\theta_a+l} \sum_{\theta_j=\theta_a-l}^{\theta_j=\theta_a+l} P_{cc}(\theta_i, \theta_j) \quad (5.4)$$

Le degré de connexité couleur. Le degré de connexité couleur est défini dans [Fontaine, 2001] comme le rapport de la probabilité de second ordre d'un ensemble de pixels par sa probabilité du premier ordre. Ainsi, le degré de connexité de l'intervalle de couleurs $[\theta_a - l, \theta_a + l]$, calculé dans la cellule $TdC(\theta_a, l)$ est donné par :

$$TdC(\theta_a, l) = \frac{P_2([\theta_a - l, \theta_a + l])}{P_1([\theta_a - l, \theta_a + l])} \quad (5.5)$$

ce qui, à partir des équations (5.2) et (5.4), donne :

$$Tdc(\theta_a, l) = \frac{\sum_{\theta_i=\theta_a-l}^{\theta_i=\theta_a+l} \sum_{\theta_j=\theta_a-l}^{\theta_j=\theta_a+l} P_{cc}(\theta_i, \theta_j)}{\sum_{\theta_i=\theta_a-l}^{\theta_i=\theta_a+l} P_1(\theta_i)} \quad (5.6)$$

Le lecteur pourra se référer à la figure 5.8a, qui montre le triangle de connexité Tdc calculé sur l'image 5.1a, le niveau de gris étant proportionnel au degré de connexité des pixels. Les cellules de faible connexité correspondent à des intervalles de couleurs ne formant pas de région dans l'image, tandis qu'une cellule de fort degré de connexité est susceptible de correspondre aux intervalles de couleurs associés à des classes significatives dans l'image. Par conséquent, le Tdc est seuillé (voir figure 5.8b) de manière à éliminer les cellules de degré de connexité non significatif. Ensuite, les classes sont extraites par recherche des maxima locaux des degrés de connexité. Il est montré par un exemple dans [Fontaine, 2001], que les maxima locaux de la ligne l du Tdc doivent être recherchés sur une fenêtre de taille $2 \times l + 1$, afin que deux maxima locaux voisins correspondent effectivement à deux ensembles de pixels distincts. Le filtrage du Tdc permet ainsi de détecter les cellules de degré de connexité qui coïncident avec un intervalle de couleurs associé à une classe dans l'image. Chaque ensemble de cellules de maxima locaux connexes est appelé *signature* de la classe. La figure 5.8c représente les signatures des classes détectées après filtrage à partir de l'image de la figure 5.1a.

Une analyse multi-échelle de chaque signature permet ensuite la détection de l'intervalle de couleurs associé à chaque classe. À cet effet, la cellule de degré de connexité maximal, de coordonnées (θ_a, l) , est détectée pour chaque signature. L'intervalle de couleurs $[\theta_a - l, \theta_a + l]$ correspond alors aux pixels constitutifs de la classe. La détermination de la classe de l'identifiant est ensuite facilitée par la connaissance des caractéristiques colorimétriques de l'encre utilisée, la classe élue étant celle dont la moyenne des pixels θ est la plus similaire à la composante couleur θ théorique. La figure 5.9 montre un exemple des classes détectées à partir de l'image 5.1a. Quatre classes ont pu être extraites, chacune correspondant approximativement à une région de l'image : couenne, tampon vétérinaire, bavures, marquage.

5.3.2 Reconnaissance du marquage par analyse de texture

La connexité des pixels associés à l'identifiant n'est pas toujours assurée, ce qui nous amène à procéder à un filtrage spatial de l'identifiant à l'aide d'outils de morphologie mathématique. Ainsi une dilatation de l'image associée à la classe du marquage par un élément structurant carré de taille 15×15 permet l'obtention d'éléments connexes significatifs, comme l'atteste la figure 5.10a. Appelons \mathcal{M}_d l'image des masques dilatée. Nous proposons alors d'introduire des informations de texture, calculées sur la composante achromatique V . L'utilisation du plan image V permet en effet d'obtenir le meilleur contraste de l'identifiant vis-à-vis de la couenne, et est donc appropriée à l'analyse de texture.

L'identifiant à détecter, de couleur très proche de celle de la chair, des hématomes ou encore des taches de graisse, possède par contre une texture propre, du fait de la présence des

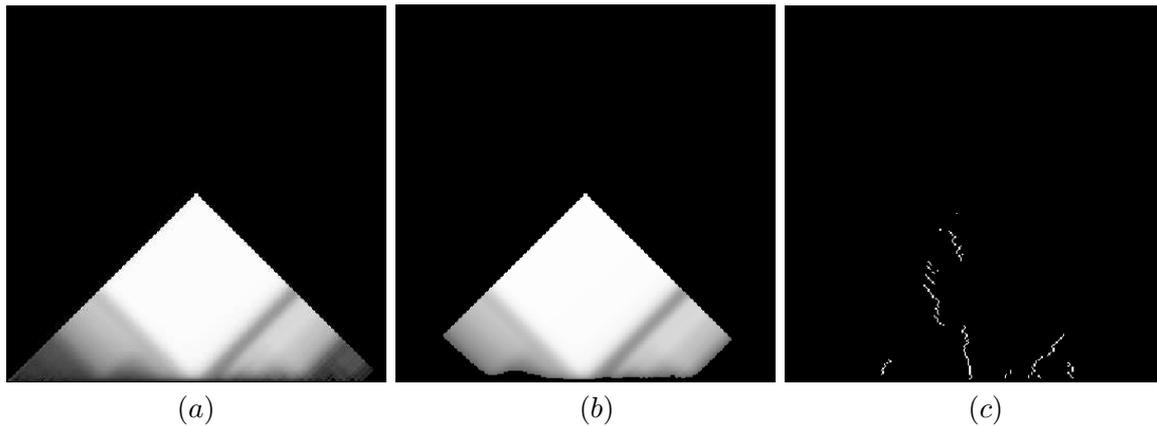


FIG. 5.8 – (a) Triangle de connexité (TdC) de la figure 5.1a. (b) TdC seuillé. (c) TdC filtré : signatures des classes.

caractères numériques. Dans un premier temps, l'image des dérivées secondes est calculée par filtrage optimal. Les filtres de Shen et Castan [Shen et Castan, 1992] sont exploités à cet effet. Leur utilisation nécessite le réglage d'un coefficient de lissage compris entre 0 et 1. Nous choisissons typiquement 0.5. Deux convolutions successives de l'image par le filtre de Shen et Castan permettent d'obtenir l'image des dérivées secondes $\nabla^2 V(p)$. La somme des valeurs absolues des dérivées secondes, calculée sur chaque composante connexe de l'image des masques \mathcal{M}_d sert de critère \mathcal{C} de détection de l'identifiant :

$$\mathcal{C} = \frac{\sum_{p \in \mathcal{M}_d} |\nabla^2 V(p)|}{N} \quad (5.7)$$

où N désigne la taille de la composante connexe sur laquelle est calculé le critère.

Ainsi, l'approche de segmentation proposée ici est dédiée à l'application de détection de l'identifiant, puisque les composantes couleur ont été sélectionnées en fonction des images à traiter. Dans un premier temps une classification des pixels est accomplie par analyse d'une structure de données multi-échelle appelée *triangle de connexité*. Par ce biais, le processus de classification intègre des informations spatiales des pixels. Enfin, l'analyse d'un critère textural finalise la détection de l'emplacement de l'identifiant. Dans [Gouiffès *et al.*, 2004c], nous avons montré la bonne robustesse de cette approche, en comparaison à une autre technique proposée, qui n'exploitait pas les attributs de texture.

La section suivante montre quelques résultats visant notamment à comparer les approches proposées.

5.4 Résultats expérimentaux

Afin de comparer les approches de segmentation, nous analysons les résultats obtenus selon le type de défaut présent dans l'image. Par souci de lisibilité, nous distinguons les différents défauts (et également l'absence de défaut) de la manière suivante :

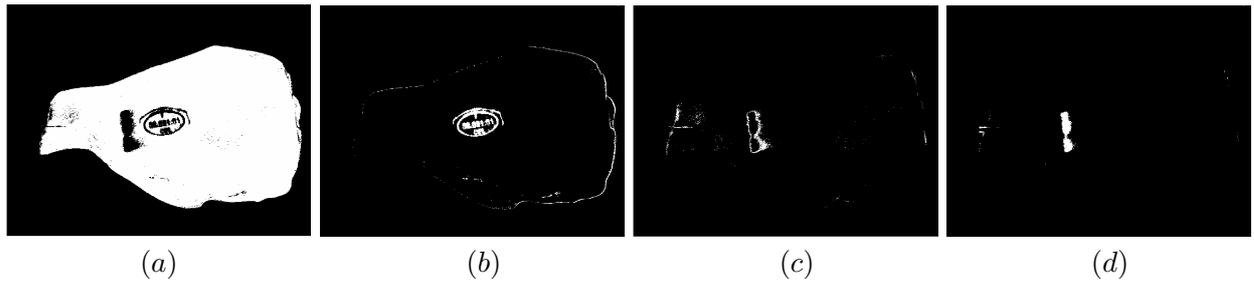


FIG. 5.9 – Classes détectées par la méthode A. (a) Couenne. (b) Tampon vétérinaire. (c) Bavure. (d) Marquage.

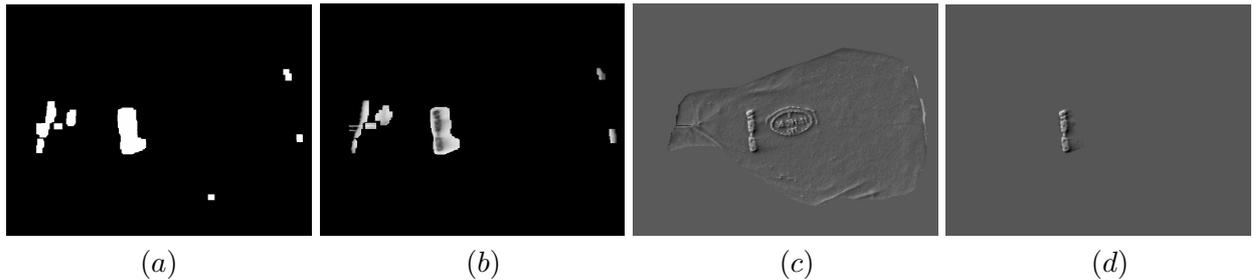


FIG. 5.10 – (a) Classe de l'identifiant filtrée par dilatation mathématique à l'aide d'un élément structurant carré de taille 15×15 . (b) Image des valeurs V associées à la classe de l'identifiant. (c) Image des dérivées secondes associées à la classe. (d) Résultat de la détection de l'identifiant par la méthode B.

- d_1 : absence de défauts
- d_2 : griffures et brûlures
- d_3 : superposition du tampon vétérinaire et de l'identifiant
- d_4 : hématomes
- d_5 : érythèmes
- d_6 : poils
- d_7 : taches de graisse
- d_8 : présence d'un second identifiant raté
- d_9 : occultation de plus de la moitié de l'identifiant

Les figures 5.11 et 5.12 montrent quelques résultats de détection de l'identifiant sur des jambons possédant des défauts. La première colonne correspond aux images initiales. Les colonnes I et II correspondent aux résultats issus de l'utilisation des attributs invariants, respectivement avec analyse et sans analyse de voisinage. Enfin, la colonne III montre les résultats obtenus avec l'approche développée dans l'espace HSV . Dans la majorité des cas, toutes ces approches assurent une détection correcte de l'emplacement de l'identifiant, quels que soient les défauts présents dans l'image. Dans le cas où l'identifiant est en partie occulté, et qu'un deuxième identifiant raté est visible, les méthodes échouent. C'est par exemple le cas des résultats de la figure 5.11d. Par contre, lorsque l'analyse de voisinage n'est pas utilisée (méthode II), les risques d'ambiguïté sont plus importants, comme le montrent les résultats 5.11-IIc et 5.12-IIb. L'algorithme a détecté soit la chair soit l'hématome au lieu de l'identifiant.

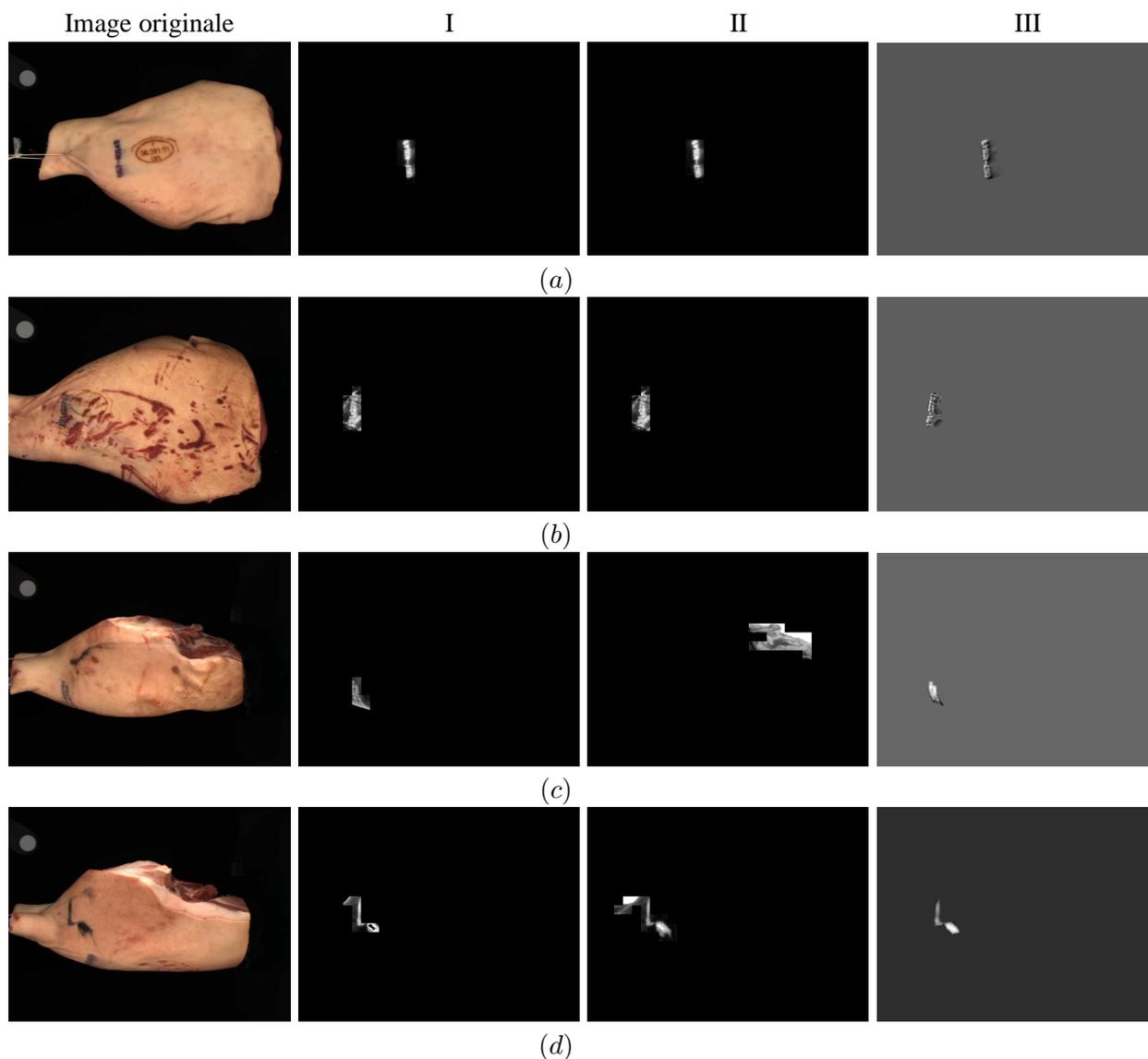


FIG. 5.11 – Résultats de segmentation. I : résultats avec utilisation d'attributs invariants dédiés à l'encre et analyse du voisinage. II : résultats avec utilisation des attributs invariants dédiés à l'encre sans analyse du voisinage. III : résultats obtenus avec l'approche de segmentation dans l'espace HSV.

Le tableau 5.2 permet d'appréhender les points forts et les faiblesses de chacune des méthodes en fonction du type de défauts présent sur la couenne. Près de 200 images de faces externes de jambon ont été classées suivant les types de défauts présents.

La deuxième colonne du tableau correspond au nombre d'images présentant un défaut d_i . Pour chacune des méthodes (I, II et III), on montre le nombre d'occurrences de fausses détections (E.) ainsi que le taux de bonnes détections associé (Tx.) et enfin le nombre d'imprécisions de détection (Nb. I.). Une imprécision est décelée dès que la détection de l'emplacement de l'identifiant s'accompagne de celle d'une seconde marque.

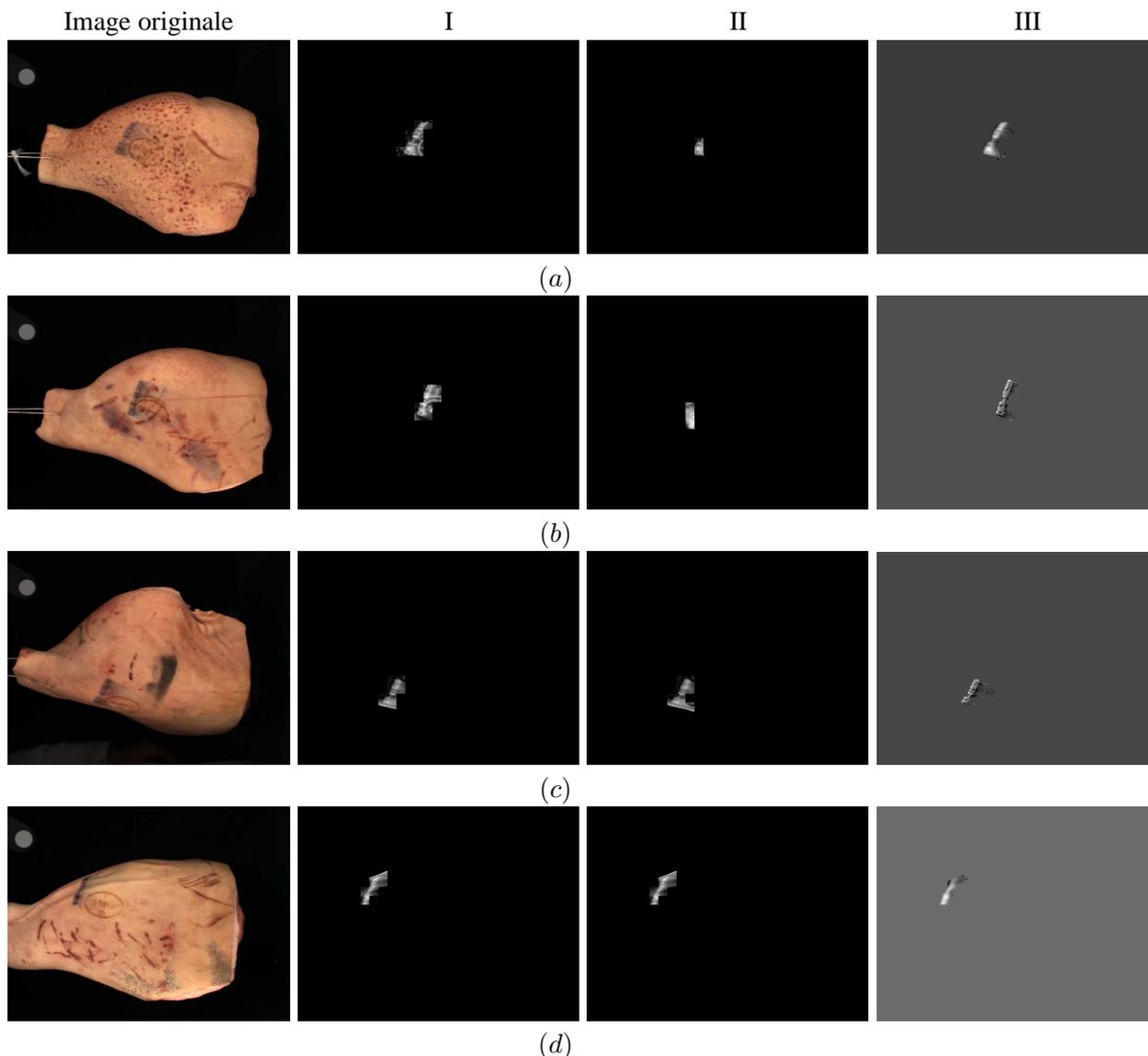


FIG. 5.12 – Résultats de segmentation (suite). I : résultats avec utilisation d'attributs invariants dédiés à l'encre et analyse du voisinage. II : résultats avec utilisation des attributs invariants dédiés à l'encre sans analyse du voisinage. III : résultats obtenus avec l'approche de segmentation dans l'espace HSV.

D'après ces résultats, les méthodes de détection I et III s'avèrent approximativement équivalentes quant à la robustesse moyenne de détection (98% pour I et 97,4 % de bonnes détections pour III). Par contre, la méthode I réalise globalement une détection plus précise. De plus, puisqu'en abattant le taux de jambons présentant des défauts est d'environ 5%, le taux de réussite global de ces méthodes est très satisfaisant.

La méthode II, qui n'effectue pas l'analyse de voisinage, obtient de moins bons résultats de segmentation. Un nombre important de fausses détections sont produites sur des jambons sans défaut et des jambons avec griffures, où les pixels d'ombres sont parfois détectés au lieu du marquage.

TAB. 5.2 – Résultats de détection obtenus par les méthodes de segmentation en fonction du type de défaut d_i . La deuxième colonne regroupe le nombre d'images possédant ce type de défaut.

Défaut	Nombre d'images	Méthode I			Méthode II			Méthode III		
		E.	Tx.(%)	Nb. I.	E.	Tx.(%)	Nb. I.	E.	Tx.(%)	Nb. I.
d_1	40	0	100	0	5	87,5	0	0	100	0
d_2	132	0	100	0	5	96,2	3	0	95,4	2
d_3	10	0	100	0	0	100	0	0	100	3
d_4	23	1	95,7	0	0	100	0	2	91,3	3
d_5	16	0	100	0	0	100	0	1	100	1
d_6	4	0	100	0	0	100	0	0	100	0
d_7	19	0	100	1	0	100	2	0	100	1
d_8	7	3	57,1	1	1	85,7	0	1	85,7	1
d_9	2	1	50	0	1	50	0	2	0	0

Méthode I : Méthode utilisant les invariants dédiés à l'encre, avec analyse de voisinage.

Méthode II : Méthode utilisant les invariants dédiés à l'encre, sans analyse de voisinage.

Méthode III : Méthode développée dans l'espace HSV .

E. : Erreur de positionnement.

Tx.(%) : Taux de détections correctes, en pourcentage.

Nb. I. : Nombre d'imprécisions de détection.

Ces résultats justifient l'utilisation de la méthode I dans le cadre de la traçabilité de la viande porcine. D'une part, elle obtient de meilleurs résultats de détection que l'approche III, qui n'effectue pas d'analyse de voisinage. D'autre part, elle s'avère plus adaptative à différents types de marquages que ne l'est la méthode II. Pour cette dernière approche, la composante couleur a été choisie en fonction des images, et un critère additionnel, autre que la couleur, est requis pour assurer une bonne détection. Or, on peut imaginer que la texture et la couleur du marquage soient amenées à être modifiées dans un abattoir. En outre, les couleurs de l'encre peuvent être différentes d'un abattoir à l'autre.

La figure 5.13 montre finalement quelques résultats de détection obtenus avec l'approche B. Les résultats de la détection sont tracés en jaune-vert. La région associée au numéro d'abattage est détectée dans des cas de figure variés, sur des jambons de teintes variables, présentant des défauts différents, et avec des opacités d'encre variables.

5.5 Conclusion

Une application de contrôle de la traçabilité de produits porcins a été présentée dans ce chapitre. Il s'agit de détecter sur les jambons la position d'un numéro d'abattage imprimé à l'encre. Dans un premier temps, les approches de segmentation basées sur l'utilisation des attributs invariants dédiés à l'encre (présentés dans le chapitre 4), ont été appliquées à ce cadre d'étude. Deux types de techniques ont ensuite été proposées. Elles se différencient entre elles par l'utilisation ou non d'une analyse spatiale.

Nous comparons ces techniques à une approche développée dans l'espace couleur *HSV*. Cette dernière se base sur une classification spatio-colorimétrique. De manière à réduire les temps de calcul, les traitements sont effectués sur un attribut issu d'une combinaison linéaire des composantes de l'espace *HSV*, ce qui permet une bonne séparabilité entre les couleurs du marquage et les autres couleurs présentes dans l'image (associées notamment aux défauts). Par contre, elle requiert l'utilisation d'un critère de texture pour détecter efficacement l'identifiant.

D'après les résultats de segmentation obtenus, la méthode alliant l'utilisation des invariants et l'analyse spatiale s'avère la plus pertinente. Elle obtient de bons résultats de détection, sans avoir à exploiter des attributs de texture. De ce fait, elle est plus facilement adaptable à d'autres formes de marquages.

Dans l'introduction, nous avons noté le fait que la couleur de l'encre utilisée pour marquer le numéro d'abattage pouvait être différente d'un abattoir à l'autre. Or les approches de segmentation basées sur les invariants dédiés à l'encre sont très peu dépendantes de la couleur de l'encre utilisée. Les paramètres à ajuster dépendent surtout du degré de transparence toléré.

Par contre, dans le cas de l'approche développée dans l'espace *HSV*, les composantes couleur ont été sélectionnées de manière à discriminer au mieux le marquage de couleur noire-bleutée des autres marques présentes sur la couenne. Or on ne peut pas affirmer que la composante utilisée permette une bonne détection quelle que soit la couleur d'encre.

À l'issue de la segmentation, le masque associé à l'identifiant doit servir de zone d'intérêt pour effectuer un déplacement de la caméra par asservissement visuel en face de l'identifiant. Après ce déplacement, une seconde étape de segmentation est nécessaire pour extraire les caractères, de façon à ce qu'ils soient reconnus par un algorithme de reconnaissance de caractères. La figure 5.14 montre quelques exemples d'images à traiter. Notons là encore la variabilité de la couleur de couleur ainsi que les différences de concentration. Dans ce cadre applicatif, nous avons également montré dans [Roué, 2004] la pertinence des attributs invariants dédiés à l'encre par rapport aux coordonnées *RGB*.

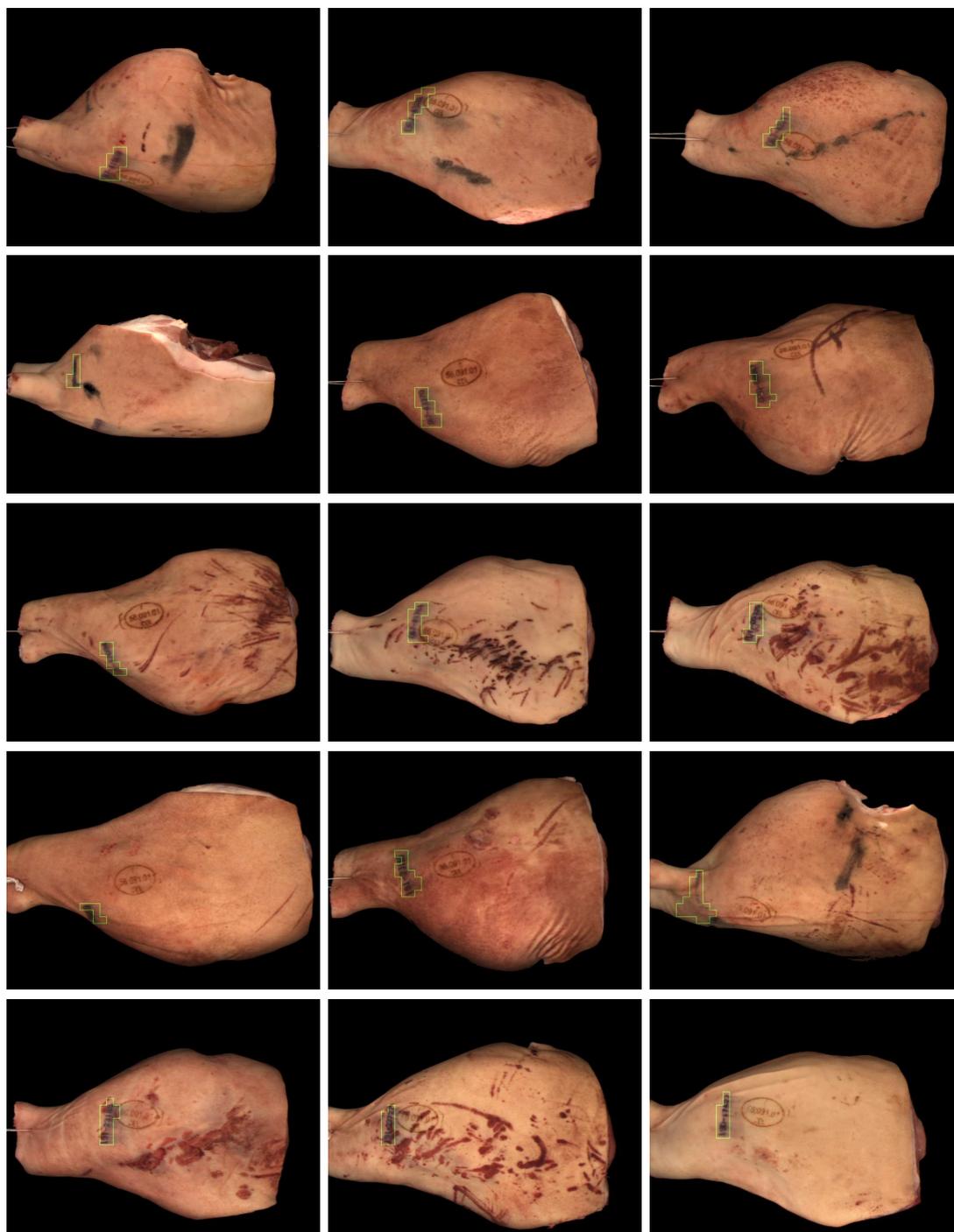


FIG. 5.13 – Exemples de résultats de détection de l'identifiant.



FIG. 5.14 – Exemples d'images de numéro d'abattage.

CONCLUSION

Dans cette partie, nous nous sommes intéressés à l'utilisation de la couleur pour une tâche de détection de primitive. Un état de l'art sur la segmentation couleur a montré la grande diversité des types de méthodes, qui se basent soit sur l'extraction de discontinuités, soit sur la recherche d'homogénéités dans l'espace colorimétrique ou dans l'espace géométrique. Dans le cas d'une détection d'une primitive de forme inconnue, mais dont la couleur est approximativement connue, les approches mixtes utilisant une classification puis une croissance de région nous paraissent les plus adaptées. La classification sert alors à sélectionner un germe permettant d'effectuer la croissance de région.

Nous avons étudié de manière plus approfondie le cas de la détection d'un marquage à l'encre, dont les variations d'opacité laissent plus ou moins transparaître le fond sur lequel il est imprimé. Des attributs invariants dédiés à la transparence de l'encre ont été définis : les *coefficients de concentration* et les *rapports de coefficients de concentration*. Les premiers correspondent à une mesure de l'opacité. Ils sont invariants vis-à-vis de la couleur du fond sur lequel est imprimée l'encre. Les seconds sont invariants par rapport la quantité d'encre appliquée sur le support ainsi qu'aux couleurs de l'encre et du support. Ces attributs ont été exploités dans le cadre d'une segmentation de marquages à l'encre de forme inconnue. D'après les résultats expérimentaux obtenus, l'algorithme de segmentation s'avère robuste malgré des variations d'opacité de l'encre au sein d'un même marquage et malgré les variations de la couleur du support sur lequel il est effectué.

Enfin, nous nous sommes penchés sur un problème de contrôle de la traçabilité des produits porcins, où il s'agissait de détecter une série de caractères imprimés à l'encre sur des pièces de jambon. Du fait de la grande variété des couleurs de couenne et de la variabilité de la quantité d'encre utilisée pour effectuer le marquage, les méthodes de segmentation basées sur les attributs invariants proposés s'avère tout particulièrement appropriées. Afin de prouver la pertinence de ces approches, nous les avons comparées à une technique de segmentation développée dans un espace *HSV*. Tout en apportant une bonne robustesse de la détection, les techniques basées sur les invariants permettent à terme une adaptivité vis-à-vis d'autres couleurs d'encres et d'autres formes d'identifiant. Après détection de la zone d'intérêt associée à l'identifiant, elle doit servir de primitive pour assurer le positionnement de la caméra par asservissement visuel. Or, cette tâche peut être entravée par l'apparition de spécularités et les variations d'illumination susceptibles d'apparaître au cours du mouvement de la caméra par rapport à l'objet.

Ainsi, alors que cette partie s'intéressait à l'utilisation de la couleur pour la détection de primitives, la prochaine section va se focaliser sur l'apport des modèles de réflexion et de la couleur pour assurer la robustesse du suivi de primitives vis-à-vis des variations d'illumination.

Troisième partie

**Apport de la couleur
et des modèles de réflexion
pour le suivi
de primitives**

INTRODUCTION

En vision par ordinateur, le problème de l'appariement ou de la mise en correspondance de primitives entre deux images reste ouvert. Certaines tâches de plus haut niveau peuvent en dépendre directement, par exemple le calcul du mouvement 2D, le suivi d'objets ou de cibles, la vision dynamique, active, ou encore l'asservissement visuel. Ce problème devient plus complexe encore dès lors que des amers visuels ne peuvent être utilisés. Citons par exemple le cas d'applications dans les domaines de la chirurgie, de l'agro-alimentaire [Kermad *et al.*, 2002], de l'agriculture, ou celles intervenant en environnement inconnu comme les applications sous marines [Tommasini *et al.*, 1998a, Espiau *et al.*, 2002] ou spatiales. Dans ce contexte, les points d'intérêt, ou encore des régions de l'image sont parfois les seules primitives extractibles.

Les primitives de type points d'intérêt correspondent à des discontinuités de reflectance⁴. Ils représentent de l'information caractéristique, bien distincte par rapport à leurs pixels voisins, en comparaison à des zones homogènes ou à des changements unidimensionnels du signal, comme les lignes de contours. En outre, leur caractère local peut représenter un atout : de par la multiplicité des points détectés dans une image, en comparaison à une détection de contours, il est peu probable que l'ensemble des primitives soit occulté au même moment.

La plupart des applications qui requièrent des mises en correspondance entre deux ou plusieurs images se basent sur la modélisation de déformations géométriques inter-images, mais négligent bien souvent les variations photométriques. Ces changements apparaissent pourtant très régulièrement lorsque les séquences d'images mises en jeu sont acquises sous un éclairage non contrôlé, ou lorsque les objets sont non lambertiens. Ce constat a également été effectué dans [Alhaj, 2004], où il s'agit d'effectuer une tâche de positionnement par asservissement visuel. La prise en compte des changements d'illumination dans cette tâche doit permettre d'en améliorer la robustesse. Cela est d'autant plus vrai dans l'application de traçabilité à laquelle nous nous intéressons, qui met en jeu un objet agro-alimentaire pouvant entraîner des réflexions spéculaires : le jambon (voir l'annexe A). Après détection de l'identifiant du jambon, telle qu'elle a été effectuée dans le chapitre 5, la motivation sous-jacente de cette partie consiste à améliorer la robustesse du positionnement de la caméra en face de cet identifiant, en compensant les variations photométriques provoquées par le déplacement de la caméra par rapport à l'éclairage, ou par l'instabilité temporelle de l'intensité de la source.

⁴Notons qu'ils peuvent correspondre également à des discontinuités de profondeur, mais ces cas de figure ne sont pas favorables du fait des occultations partielles

Les variations d'illumination peuvent être modélisées localement dans un voisinage autour du point à suivre, aussi bien dans les images de luminance que dans les images couleur. En outre, l'information trichromatique disponible dans les séquences d'images couleur permet d'extraire des attributs invariants vis-à-vis des changements d'illumination. Ces attributs, très fréquemment utilisés en suivi d'objets, en reconnaissance ou en segmentation, n'ont jamais été exploités dans le cadre du suivi de points. Cette partie du document s'intéresse à ces problématiques.

Premièrement, en se basant sur les modèles de réflexion les plus usités en vision par ordinateur, qui ont été présentés dans la section 1.2, nous proposons une justification des modélisations locales de changements d'illumination utilisées dans les tâches de mise en correspondance ou de suivi de primitives. Les hypothèses sur lesquelles ils se basent seront explicitées. Nous privilégions deux modèles permettant de compenser les variations spatiales de changement d'illumination. À notre connaissance, ces modèles d'illumination n'ont pas été utilisés en suivi de points différentiel. Nous étudions également la validité de ce type de modèle photométrique en fonction des paramètres de la scène : la courbure et les propriétés optiques de la surface, les positions de l'éclairage, de la caméra et de l'objet. Cette étude fait l'objet du chapitre 6.

Ensuite, les modélisations proposées sont exploitées dans le cadre du suivi de primitives de type point dans les images de luminance, où les paramètres de mouvement du point ainsi que les paramètres de variations d'illumination sont approximés tout au long de la séquence d'images. Nous comparons cette approche aux techniques de suivi de points existantes. Cette méthode est également étendue au cas du suivi de zones d'intérêt, en considérant des tailles de voisinage plus larges que pour le suivi de points d'intérêt. Le chapitre 7 est ainsi consacré au suivi dans les images de luminance.

Contrairement aux images de luminance, où les compensations d'illumination s'effectuent de manière locale dans l'image, les images couleur permettent d'extraire des attributs invariants en tout point de l'image, de manière globale. Le chapitre 8 dresse un état de l'art des différents invariants proposés dans la littérature, en fonction des hypothèses sur lesquelles ils se fondent. Ces hypothèses concernent les propriétés optiques des matériaux mais également certaines conditions d'acquisition.

Pour finir, nous nous intéressons à la robustesse du suivi de points couleur vis-à-vis des variations d'illumination. D'une part, la modélisation locale utilisée dans le cadre du suivi de points dans les images de luminance est étendue au cas de la couleur. D'autre part, les attributs invariants globaux sont exploités. Enfin, étant donné que les attributs invariants ne compensent pas parfaitement toutes les variations d'illumination potentielles, nous les allions avec la modélisation photométrique locale. Cette étude est présentée dans le chapitre 9.

MODÉLISATION DES VARIATIONS PHOTOMÉTRIQUES DANS DES SÉQUENCES D'IMAGES

Les tâches de mises en correspondance ou de suivi entre deux images d'une même scène sont sensibles aux variations d'illumination. En se basant sur les modèles de réflexion exposés dans le chapitre 1, nous proposons une modélisation des variations d'illumination susceptibles de se produire entre deux prises de vue, que ces changements soient spéculaires ou qu'ils soient dus à l'éclairage. Puisque ces modèles requièrent un nombre important de paramètres, qui ne sont pas toujours disponibles sans connaissance a priori sur les matériaux mis en jeu, seuls des modèles locaux sont facilement utilisables en vision par ordinateur. Dans ce chapitre, nous justifions analytiquement leur utilisation. En outre, nous proposons de modéliser localement, par le biais de deux formalismes, les variations spatiales de changements d'illumination. En étudiant la validité de cette modélisation, nous montrons sa pertinence par rapport aux modèles les plus généralement utilisés en vision par ordinateur. Ces modèles seront exploités dans le chapitre 7 dans le cadre du suivi de points d'intérêt et de motifs plus larges de l'image.



L'appariement et la mise en correspondance de points d'intérêt et de motifs d'intérêt entre deux prises de vue d'une même scène doivent être améliorés par la prise en compte des variations d'illumination. Premièrement, nous décrivons dans la section 6.1 une modélisation précise des différents types de changements d'illumination en se basant sur les modèles de réflexion abordés dans la section 1.2. Puisque ce formalisme exige un grand nombre de paramètres, son utilisation en vision par ordinateur s'avère compromise. Des modèles simplifiés, valables localement dans l'image, sont alors proposés. Dans la section 6.2, nous décrivons sur quelles hypothèses ils se basent. La section 6.3 propose deux modèles qui permettent de compenser au mieux les changements d'illumination. Tandis que le premier de ces formalismes ne peut être

validé que dans des zones très locales de l'image, telles que de petites fenêtres centrées sur des points d'intérêt, le second est approprié à des variations plus globales sur des motifs d'intérêt de l'image. Enfin, l'étude de la validité des modèles proposés est effectuée dans la section 6.4. Elle prend en compte la géométrie de la surface, ses propriétés optiques ainsi que les positions de l'éclairage, de la caméra et de la surface.

6.1 Variations de luminance et de couleur

Considérons un capteur, susceptible d'être en mouvement, qui acquiert une séquence d'images d'une scène dont les objets peuvent également être mobiles. En outre, la position et l'intensité de la source d'éclairage ne sont pas considérées constantes au cours du temps. Soient f et g deux images acquises respectivement aux instants k et k' . Après un déplacement relatif entre la caméra et les objets de la scène, le point P de la scène ne se projette plus sur le capteur en p de coordonnées (x_p, y_p) mais en p' de coordonnées (x'_p, y'_p) . La figure 6.1 décrit les vecteurs et les angles utilisés dans ce chapitre. Comme dans le chapitre 1, \mathbf{V} et \mathbf{L} sont respectivement les directions d'observation et d'éclairage, qui forment des angles θ_r et θ_i avec la normale \mathbf{n} en P . \mathbf{B} est la bissectrice de l'angle entre \mathbf{V} et \mathbf{L} , formant un angle ρ avec la normale \mathbf{n} .

Dans un premier temps, nous considérons le cas des changements de luminance et de couleur dus à la réflexion spéculaire à la surface des objets. Le cas des variations de luminance dues conjointement aux spécularités et aux variations d'éclairage est ensuite étudié. Notons que les changements de luminance dus aux problèmes liés à l'acquisition (distorsions dues à l'objectif, flou) ne sont pas pris en compte dans cette étude.

La luminance a été mise en relation avec l'irradiance sur le capteur dans la section 2.1.1. Puisque $\mathcal{L}(\lambda, P) = \mathcal{E}(\lambda, P)\mathcal{R}(\lambda, P)$, l'expression de $f(p)$ est ainsi donnée par :

$$f(p) = K_c \int_{\lambda_{\min}}^{\lambda_{\max}} \mathcal{S}(\lambda)\mathcal{E}(\lambda, P)\mathcal{R}(\lambda, P)d\lambda \quad (6.1)$$

où $\mathcal{R}(\lambda, P)$ est la reflectance du matériau, $\mathcal{S}(\lambda)$ est la courbe de sensibilité du capteur et $\mathcal{E}(\lambda, P)$ la forme du spectre associé à l'illuminant. Posons $\mathcal{M}(\lambda) = \mathcal{S}(\lambda)\mathcal{E}(\lambda, P)$. Dans le cas d'un capteur à réponse linéaire et d'un illuminant dont la couleur n'est pas modifiée, le terme $\mathcal{M}(\lambda)$ peut être exprimé comme le produit d'un gain K_m indépendant des longueurs d'ondes avec un spectre de valeur maximale unitaire, que l'on appelle $e(\lambda)$. Dans ce cas :

$$f(p) = K_c K_m(p) \int_{\lambda_{\min}}^{\lambda_{\max}} e(\lambda)\mathcal{R}(\lambda, P)d\lambda \quad (6.2)$$

D'après les modèles de réflexion décrits dans la section 1.2.2, la reflectance s'exprime par la somme d'un terme dépendant de la reflectance diffuse $\mathcal{R}_b(\lambda, P)$ et d'un terme spéculaire dépendant de l'illuminant. Notons $a(p)$ le terme exprimé ci-dessous par :

$$a(p) = \int_{\lambda_{\min}}^{\lambda_{\max}} e(\lambda)\mathcal{R}_b(\lambda, P)d\lambda \quad (6.3)$$

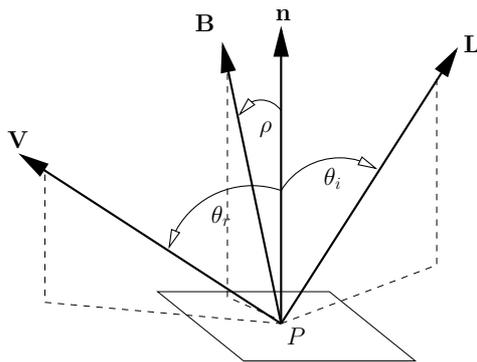


FIG. 6.1 – Vecteurs et angles utilisés pour décrire les réflexions.

où \mathcal{R}_b correspond à la reflectance diffuse du matériau.

Quel que soit le modèle de réflexion utilisé, les intensités de l'image f peuvent être modélisées par la somme de trois termes reliés respectivement aux réflexions diffuse, spéculaire et ambiante :

$$f(p) = K_d(p)a(p) \cos \theta_i(P) + K_s(p)h_f(P) + K_a(p) \quad (6.4)$$

où $K_d(p)$, $K_s(p)$ et $K_a(p)$ dépendent à la fois du gain de la caméra et de l'intensité d'éclairage, puisqu'ils sont proportionnels à $K_c K_m$. Rappelons que, dans le cas du modèle de Phong, la fonction de spécularité h_f s'écrit :

$$h_f(P) = \cos^n(\rho(P)) \quad (6.5)$$

et dans le cas d'un modèle de Torrance-Sparrow :

$$h_f(p) = \frac{e^{-\frac{\rho^2(P)}{2\sigma^2}}}{\cos(\theta_r(P))} \quad (6.6)$$

L'information couleur est quant à elle obtenue après intégration de l'irradiance sur chacun des trois capteurs, avec des gains qui peuvent être différents. La couleur \mathbf{f} peut alors s'écrire de manière vectorielle :

$$\mathbf{f}(p) = \mathbf{C}_d(p) \cdot \mathbf{a}(p) \cos \theta_i(P) + \mathbf{C}_s(p)h_f(P) + \mathbf{C}_a(p) \quad (6.7)$$

où « \cdot » est le produit d'Hadamard. Les termes $\mathbf{C}_d(p)$, $\mathbf{C}_s(p)$ et $\mathbf{C}_a(p)$ ne dépendent que de la couleur de l'illuminant et des gains des capteurs. Quant à $\mathbf{a}(p)$, il dépend de la reflectance diffuse du matériau.

Voyons à présent comment se caractérisent les variations de spécularité apparues entre deux images de luminance f et g , puis entre deux images couleur \mathbf{f} et \mathbf{g} . Nous notons K_a' , K_d' , K_s' , a' , θ_i' et h_g les termes associés à l'image de luminance g , et \mathbf{C}_a' , \mathbf{C}_d' , \mathbf{C}_s' , \mathbf{a} ceux associés à l'image couleur \mathbf{g} .

6.1.1 Variations de spécularité

Afin de ne prendre en compte que l'effet des spécularités, nous considérons dans un premier temps le cas d'une caméra en mouvement observant un objet immobile.

Images de luminance. En supposant une unique source d'éclairage ponctuelle et immobile (c'est-à-dire que $\theta_i'(P) = \theta_i(P)$), d'irradiance constante au cours du temps ($K_d(p')a(p') = K_d(p)a(p)$) la réflexion diffuse $K_d(p)a(p) \cos \theta_i(P)$ ne varie pas. Nous admettons également que les termes d'éclairage ambiant K_a et direct K_s sont constants au cours du temps. La composante spéculaire est quant à elle susceptible de changer de manière très significative au cours du déplacement de la caméra vis-à-vis de l'objet. La luminance g en p' à l'instant k' est alors exprimée par :

$$g(p') = K_d(p)a(p) \cos \theta_i(P) + K_s(p)h_g(P) + K_a(p) \quad (6.8)$$

où $h_g(P)$ correspond au terme de réflexion spéculaire, qui peut être décrit par exemple par le modèle de Phong (6.5) ou le modèle de Torrance-Sparrow (6.6). Ainsi, la relation entre les images f et g à deux instants k et k' est donnée par :

$$g(p') = f(p) + K_s(p)[h_g(P) - h_f(P)] \quad (6.9)$$

La relation générale (6.9) reste valable dans le cas de plusieurs sources d'éclairage, mais $h_g(P)$ et $h_f(P)$ sont alors associés à la somme de plusieurs composantes spéculaires, théoriquement autant de termes qu'il y a de vecteurs d'éclairage. Dans la suite, nous posons :

$$\psi(p) = K_s(p)[h_g(P) - h_f(P)] \quad (6.10)$$

Il n'existe ainsi qu'une façon d'exprimer les variations de spécularité entre f et g :

$$\boxed{g(p') = f(p) + \psi(p)} \quad (6.11)$$

Images couleur. Dans le cas des images couleur, la couleur $\mathbf{g}(p')$ au point p' s'exprime de manière vectorielle :

$$\mathbf{g}(p') = \mathbf{C}_a(p) + \mathbf{C}_d(p) \cdot \mathbf{a}(p) \cos \theta_i(P) + \mathbf{C}_s(p)h_g(P) \quad (6.12)$$

Là encore, l'intensité d'éclairage est supposée constante, de sorte que les différentes couleurs mises en jeu \mathbf{C}_a , \mathbf{C}_d et \mathbf{C}_s demeurent constantes. Ainsi, les images \mathbf{g} et \mathbf{f} sont reliées de la même manière que (6.11) :

$$\boxed{\mathbf{g}(p') = \mathbf{f}(p) + \boldsymbol{\psi}(p)} \quad (6.13)$$

La fonction $\boldsymbol{\psi}(p)$ est définie de façon similaire à (6.10), en remplaçant toutefois K_s par \mathbf{C}_s :

$$\boldsymbol{\psi}(p) = \mathbf{C}_s(p)[h_g(P) - h_f(P)] \quad (6.14)$$

Notons que dans le cas d'un illuminant blanc (et de gains identiques sur R , G et B) $\mathbf{C}_s = (1, 1, 1)^T$, ce qui revient à considérer la fonction $\boldsymbol{\psi}$ comme une fonction scalaire, comme le décrit (6.10).

La prochaine section s'intéresse aux changements de luminance dus à la fois aux variations d'éclairage et de spécularité.

6.1.2 Variations d'éclairage et de spécularité

Images de luminance. Un changement d'intensité de l'éclairage provoqué entre deux luminances $f(p)$ et $g(p')$ associées à un même point P à deux instants k et k' peut intervenir de trois manières différentes :

1. une variation d'intensité t_a de l'éclairage uniforme K_a ou, ce qui est équivalent, une modification du gain de la caméra. Le terme d'éclairage ambiant $K'_a(p')$ en p' dans l'image g devient alors $K'_a(p') = K_a(p) + t_a(p)$;
2. une variation d'intensité t_d de l'éclairage direct K_d et de l'intensité de spécularité K_s (ou d'un changement de gain de la caméra) : $K'_d(p') = K_d(p) + t_d(p)$ et $K'_s(p') = K_s(p) + t_s(p)$;
3. un déplacement de la source directe ou de l'objet, qui entraîne des variations $t_i(P)$, $t_r(P)$ et $t_\rho(P)$ des angles $\theta_i(P)$, $\theta_r(P)$ et $\rho(P)$, tel que $\theta'_i(P) = \theta_i(P) + t_i(P)$, $\theta'_r(P) = \theta_r(P) + t_r(P)$ et $\rho'(P) = \rho(P) + t_\rho(P)$

Dans le cas où le spectre de l'illuminant ne varie pas (la couleur de l'illuminant reste identique), $a(p)$ dépend uniquement de la reflectance et est donc constant au cours du temps. En intégrant les variations d'illumination dans l'expression de la luminance (6.4), l'image de luminance à l'instant k' s'exprime :

$$g(p') = K'_d(p')a(p) \cos \theta'_i(P) + K'_s(p')h_g(P) + K'_a(p') \quad (6.15)$$

En prenant en compte l'expression des variations d'illumination, cette formulation devient :

$$g(p') = (K_d(p) + t_d(p))a(p) \cos(\theta_i(P) + t_i(P)) + (K_s(p) + t_s(P))h_g(P) + K_a(p) + t_a(p) \quad (6.16)$$

La fonction $h_g(P)$ comprend les changements d'angle $\rho(P)$ et $\cos \theta_r$ dus au déplacement de la source d'éclairage direct ou au déplacement de la caméra. Il existe deux manières de modéliser ces variations.

◦ Premièrement, en développant $\cos(\theta_i(P) + t_i(P))$ dans l'équation (6.16), on y fait apparaître l'expression de $f(p)$ décrite par (6.4). La luminance $g(p')$ s'exprime donc en fonction de $f(p)$ de la façon suivante :

$$\boxed{g(p') = f(p) + \phi(p)} \quad (6.17)$$

où la fonction $\phi(p)$ est donnée, tous calculs faits, par :

$$\phi(p) = t_a(p) + t_d(p)a(p)[\cos \theta_i(P) \cos t_i(P) - \sin \theta_i(P) \sin t_i(P)] + t_s(p)h_g(P) + \psi(p) \quad (6.18)$$

Lorsque l'intensité de l'éclairage varie, c'est-à-dire que $t_d(p) \neq 0$ (voir l'équation (6.18)), ou lorsque la source se déplace entre k et k' (c'est-à-dire que $t_i \neq 0$), $\phi(p)$ dépend de $a(p)$, et par conséquent de la reflectance du matériau.

◦ D'une autre façon, en exprimant $a(p)$ à partir de (6.4) puis en injectant son expression dans (6.16), il vient la relation suivante entre les luminances f et g :

$$\boxed{g(p') = \lambda(p)f(p) + \eta(p)} \quad (6.19)$$

Les fonctions λ et η sont définies comme suit :

$$\lambda(p) = \frac{(K_d(p) + t_d(p)) \cos(\theta_i(P) + t_i(P))}{K_d(p) \cos \theta_i(P)} \quad (6.20)$$

$$\eta(p) = -(K_s(p)h_f(P) + K_a(p))\lambda(p) + (K_s(p) + t_s(p))h_g(P) + K_a(p) + t_a(p) \quad (6.21)$$

Contrairement à la relation (6.17), les fonctions $\lambda(p)$ et $\eta(p)$ ne modélisent pas la reflectance de la surface (elles ne dépendent pas de $a(p)$), mais dépendent uniquement des termes relatifs à la géométrie de la scène.

La formulation (6.19) peut être trouvée dans [Negahdaripour, 1998, Black *et al.*, 2000], mais une expression analytique de λ et η n'a pas été définie à notre connaissance.

Images couleur. Dans le cas des images couleur, trois types de variations peuvent également être provoquées :

1. une variation de couleur t_a de l'éclairage ambiant C_a ou, ce qui est équivalent, une modification du gain de la caméra sur les trois canaux R, G et B : $C'_a(p') = C_a(p) + t_a(p)$;
2. des variations $t_i(P), t_r(P)$ et $t_\rho(P)$ des angles $\theta_i(P), \theta_r(P)$ et de $\rho(P)$;
3. des variations de couleur t_d et t_s de l'éclairage direct C_d et de la couleur de spécularité C_s (ou un changement de gain de la caméra) : $C'_d(p') = C_d(p) + t_d(p)$ et $C'_s(p') = C_s(p) + t_s(p)$.

En reprenant le même raisonnement que celui effectué pour les images de luminance, deux relations permettent d'exprimer les variations d'illumination dans les images couleur.

◦ La première les décrit à l'aide d'une seule fonction ϕ :

$$\boxed{g(p') = f(p) + \phi(p)} \quad (6.22)$$

où ϕ dépend de la reflectance par le biais de a :

$$\phi(p) = t_a(p) + t_d(p) \cdot a(p) [\cos \theta_i(P) \cos t_i(P) - \sin \theta_i(P) \sin t_i(P)] + t_s(p)h_g(P) + \psi(p) \quad (6.23)$$

◦ La seconde généralise la relation (6.19) au cas des images de couleur, en obtenant un second modèle, ne dépendant pas de la reflectance :

$$\boxed{g(p') = \lambda(p) \cdot f(p) + \eta(p)} \quad (6.24)$$

Les différentes fonctions de changement d'illumination deviennent vectorielles de sorte que :

$$\lambda(p) = \frac{(C_d(p) + t_d(p)) \cos(\theta_i(P) + t_i(P))}{C_d(p) \cos \theta_i(P)} \quad (6.25)$$

$$\eta(p) = -(C_s(p)h_f(P) + C_a(p))\lambda(p) + (C_s(p) + t_s(p))h_g(P) + C_a(p) + t_a(p) \quad (6.26)$$

où l'opérateur de division correspond à une division terme à terme des vecteurs mis en jeu. Notons toutefois que dans le cas où l'illuminant est blanc et le demeure entre les deux instants k et k' , les valeurs C_d, C_s et C_a sont scalaires et (6.24) se simplifie de la manière suivante :

$$g(p') = \lambda(p) f(p) + \eta(p) \quad (6.27)$$

Les fonctions de variation d'illumination ne dépendent plus des variations de couleur de l'illuminant, mais uniquement de ses changements d'intensité.

Étant donné que les modèles de variation d'illumination présentés dans cette section se réfèrent à un grand nombre de paramètres, leur utilisation en vision par ordinateur n'est pas triviale. Puisqu'ils dépendent des caractéristiques des matériaux, *via* la rugosité ou encore de la forme par l'intermédiaire des normales en chaque point, ils sont difficilement estimables. Par contre, des approximations locales sont souvent intégrées dans les travaux requérant des mises en correspondance locales entre deux images : en indexation [Gros, 2000], en suivi de points [Jin *et al.*, 2001], en calcul de flot optique [Negahdaripour, 1998].

6.2 Modèles locaux de variation d'illumination

Par définition, les modèles photométriques locaux sont valables dans des zones restreintes de l'image. Dans notre cadre d'étude, nous notons \mathcal{W} cette fenêtre, qui est centrée sur un point p . On appelle m un point appartenant à \mathcal{W} , correspondant à la projection d'un point M de la scène. Nous analysons ici, à partir de (6.19), (6.20) et (6.21) (respectivement (6.24), (6.25) et (6.26) dans le cas de la couleur) sur quelles hypothèses implicites ils se fondent.

6.2.1 Modèle de conservation de la luminance

Dans bon nombre d'applications, la luminance est supposée constante au cours du temps en accord avec la contrainte de conservation de la luminance émise par [Horn et Schunck, 1981]. La relation entre deux luminances à des instants k et k' s'écrit alors :

$$g(m') = f(m) \text{ pour tout } m \in \mathcal{W} \quad (6.28)$$

Selon (6.19), cette égalité est vraie si $\lambda(m) = 1$ et $\eta(m) = 0$ et ce $\forall m \in \mathcal{W}$. D'après (6.20), la solution la plus évidente consiste à ce que les fonctions $t_d(m)$ et $t_i(M)$ soient nulles quel que soit $m \in \mathcal{W}$, c'est-à-dire lorsqu'aucun changement d'éclairage n'a lieu. Il s'agit d'une condition suffisante mais pas forcément nécessaire, puisqu'il est possible d'imaginer une configuration d'éclairage et de géométrie de surface telle qu'une combinaison de valeurs $t_d(m)$ et $t_i(M)$ satisfasse également l'égalité $\lambda(m) = 1$. Cependant, ce postulat s'avère difficilement valide simultanément en tout point dans l'image et pour tout type d'objet. Il s'agit en fait d'un cas dégénéré.

De même, en considérant $\lambda(m) = 1$ dans l'équation (6.21), la solution la plus évidente pour que le terme $\eta(m)$ soit annulé est de ne prendre en compte aucun changement d'éclairage ambiant $t_a(p) = 0$ et aucune variation de spécularité $h_f(M) = h_g(M)$, $\forall m \in \mathcal{W}$. Si l'on considère un matériau spéculaire, pour lequel h_f et h_g sont exprimés par exemple par les modèles de Phong (6.5) ou de Torrance-Sparrow (6.6), ces fonctions varient dès que le vecteur d'éclairage ou d'observation change, c'est-à-dire dès qu'un mouvement relatif objet/caméra ou objet/éclairage est effectué. Soulignons donc que le postulat $h_f(M) = h_g(M) \forall m \in \mathcal{W}$ n'est valable en tout point m que si le matériau de l'objet est lambertien, ou encore si la géométrie de la scène est fixe (ni déplacement source/objet ni déplacement caméra/objet). Là encore, il peut

exister des solutions singulières qui satisfont l'égalité $\eta(m) = 0 \forall m \in \mathcal{W}$. Par exemple, d'après l'équation (6.21), cela peut être réalisé lorsque $(K_s(m)h_f(M) + K_a(m))\lambda(m) = (K_s(m) + t_s(m))h_g(M) + K_a(m) + t_a(m)$. Il s'agit évidemment d'un cas particulier puisque cette égalité suppose forcément certaines hypothèses très singulières, notamment quant à la courbure de la surface ou la rugosité du matériau. Ainsi, l'hypothèse lambertienne et l'absence de variation d'éclairage doivent être assurées.

Examinons à présent le cas des images couleur. L'hypothèse de conservation implique que $\mathbf{g}(m) = \mathbf{f}(m)$ pour tout $m \in \mathcal{W}$, ce qui suppose d'après (6.24) que $\boldsymbol{\lambda} = (1, 1, 1)^T$ et $\boldsymbol{\eta} = (0, 0, 0)^T$. Outre les hypothèses que stipule le modèle de conservation de la luminance, le modèle de conservation de la couleur suppose que la couleur de l'illuminant ne varie pas ($\mathbf{t}_d = \mathbf{t}_s = \mathbf{t}_a = (0, 0, 0)^T$ dans (6.25) et (6.26)).

6.2.2 Le modèle additif

Le modèle additif exprime les changements d'illumination provoqués entre deux images f et g par la relation $g(m') = f(m) + \eta$ [Gros, 2000]. À l'instar du modèle de conservation, $\lambda(m) = 1 \forall m \in \mathcal{W}$. Par ailleurs, ce modèle suppose que η , donné par (6.21), est identique en tout point m de \mathcal{W} . De ce fait, les fonctions spéculaires h_g et h_f ne sont plus supposées nulles mais constantes en tout point de \mathcal{W} . Selon les modèles de réflexion (6.5) ou (6.6), cette affirmation est vraie uniquement si l'angle $\rho(M)$ et la rugosité (n pour le modèle de Phong et ς pour celui de Torrance-Sparrow) sont tous deux constants sur \mathcal{W} . Le postulat $h_f = h_g \forall m \in \mathcal{W}$ s'avère également exact dans le cas de matériaux lambertiens, pour lesquels $h_f = h_g = 0$, par définition. En ce qui concerne les images couleur, la relation $\mathbf{g}(m') = \mathbf{f}(m) + \boldsymbol{\eta}$ est exacte dans les mêmes conditions que pour les images de luminance, excepté qu'il est nécessaire de surcroît que l'illuminant ne change pas de couleur.

6.2.3 Le modèle scalaire

Le changement scalaire d'illumination [Gros, 2000] est modélisé par $g(m') = \lambda f(m)$ pour tout $m \in \mathcal{W}$. Selon (6.20) et (6.21), il est valide si λ est constant en tout point de \mathcal{W} , et si $\eta = 0$. Ce cas de figure se produit lorsqu'il n'y a aucun changement d'éclairage, que ce soit en intensité $t_d = 0$ ou en position $t_i = 0$ (dans ce cas λ est constant $\forall m \in \mathcal{W}$), lorsque l'objet est lambertien ($h_f = h_g = 0$) et qu'il n'y a pas d'éclairage ambiant ($K_a = t_a = 0$); Dans le cas d'images couleur, le modèle scalaire s'exprime $\mathbf{g}(m') = \boldsymbol{\lambda} \cdot \mathbf{f}(m)$ et suppose les mêmes hypothèses.

6.2.4 Le modèle affine

Le modèle affine stipule que les changements d'illumination et de spécularité sont constants sur toute la fenêtre \mathcal{W} de sorte que $g(m') = \lambda f(m) + \eta$. Il est utilisé dans de nombreuses publications, notamment dans des contextes de calcul de flot optique [Negahdaripour, 1998, Kim et al., 2005, Altunbasak et al., 2003] et de suivi de points d'intérêt [Jin et al., 2001].

D'après (6.20) et (6.21), ce formalisme suppose que l'ensemble des paramètres ($K_d, t_d, K_a, t_a, \theta_i, t_i, h_f, h_g$) sont constants en tout point m de \mathcal{W} . Cette hypothèse a également été effectuée dans [Mukawa, 1990] pour estimer les paramètres de la géométrie de la scène (l'estimation de la forme, la direction d'éclairage).

Dans le cas d'images couleur et d'un illuminant coloré, le modèle affine s'exprime avec des paramètres d'illumination vectoriels de sorte que $g(m) = \lambda \cdot f(m) + \eta$. D'après (6.25) et (6.26), cela suppose que l'ensemble des paramètres d'illumination $C_d, t_d, K_a, t_a, \theta_i, t_i, h_f, h_g$ soient constants entre k et k' .

Or, les termes θ_i, t_i, h_f et h_g sont liés à la géométrie de la scène et sont non-constants dès lors que la surface observée est non plane. Les fonctions spéculaires h_f et h_g dépendent en outre de la rugosité du matériau, qui est susceptible de varier sur \mathcal{W} .

L'ensemble des modèles photométriques qui viennent d'être décrits supposent implicitement que les variations d'illumination sont strictement constantes en tout point p de \mathcal{W} , ce qui n'est pas avéré dans le cas général, notamment lorsque les zones d'intérêt \mathcal{W} représentent des surfaces très incurvées ou encore des surfaces de rugosité variable. Dans la section qui suit, nous justifions de manière théorique deux modèles photométriques plus appropriés, dans le sens où ils ne requièrent pas l'hypothèse de constance des paramètres.

6.3 Modèles de réflexion locaux adaptés à l'occurrence de spéularités et aux variations d'éclairage

Les approximations des modèles d'illumination abordés précédemment se basent sur des hypothèses assez restrictives, telles que la constance de la géométrie de la scène ou la constance des variations d'illumination. Ces hypothèses peuvent s'avérer incorrectes. Nous proposons ici deux modèles autorisant la variation spatiale des changements d'illumination sur \mathcal{W} , qu'ils soient dus à des variations d'éclairage ou à l'apparition de spéularités. Tandis que le premier modèle est valable très localement dans l'image, le second permet de prendre en compte des motifs plus étendus.

6.3.1 Modèle d'illumination local

La section 6.1 a montré comment chaque type de changement d'illumination peut être exprimé à partir des modèles de réflexion. Les variations de spéularité entre f et g ne peuvent être exprimées que par ψ donnée par (6.10). Dans le cas général où spéularités et variations d'éclairage apparaissent conjointement, (6.17) correspond à une relation exacte. Selon les modèles de réflexion (voir (6.5) et (6.6)), la fonction ϕ , décrite par (6.18), est variable sur \mathcal{W} puisqu'elle dépend des angles d'observation et d'éclairage et par conséquent de la normale \mathbf{n} en chaque point de \mathcal{W} . Elle dépend également des caractéristiques du matériau notamment la rugosité de la surface. En outre, dans le cas d'un changement d'intensité d'éclairage, ou lorsqu'il y a déplacement relatif source/objet, ϕ s'exprime en fonction de la reflectance du matériau (par le biais de $a(p)$).

Nous admettons que ϕ peut être correctement approximée sur la zone \mathcal{W} par une fonction de classe C^K que nous notons ϕ_{mod} .

Images de luminance. Considérons un point m de coordonnées (x, y) , appartenant à la fenêtre d'intérêt \mathcal{W} centrée sur p , et qui est la projection d'un point M de la scène. Un développement de ϕ_{mod} en séries de Taylor à l'ordre K peut être effectué en m :

$$\phi_{mod}(m) \simeq \sum_{i,j,i+j=0}^{i+j=K} \phi_{ij}(p)(x - x_p)^i(y - y_p)^j \quad (6.29)$$

où les termes ϕ_{ij} correspondent aux dérivées partielles de la fonction ϕ_{mod} au point p , d'ordre i en x et j en y . En négligeant les termes supérieurs au premier ordre, on obtient finalement :

$$\phi_{mod}(m) \simeq \phi_{mod}(p) + \left. \frac{\partial \phi_{mod}}{\partial x} \right|_p (x - x_p) + \left. \frac{\partial \phi_{mod}}{\partial y} \right|_p (y - y_p). \quad (6.30)$$

Appelons $\alpha = \left. \frac{\partial \phi_{mod}}{\partial x} \right|_p$, $\beta = \left. \frac{\partial \phi_{mod}}{\partial y} \right|_p$ et $\gamma = \phi_{mod}(p)$. En injectant (6.30) dans (6.11) il vient :

$$\boxed{g(m') = f(m) + \alpha(x - x_p) + \beta(y - y_p) + \gamma} \quad (6.31)$$

Contrairement aux modèles d'illumination décrits précédemment, l'angle d'incidence θ_i et l'angle ρ ne sont pas considérés constants sur \mathcal{W} . Les paramètres K_s ainsi que les caractéristiques de rugosité des matériaux n (ou ς) peuvent être différents en tout point de \mathcal{W} . Par conséquent, les variations de spécularité peuvent être différentes en chaque point de \mathcal{W} .

Images couleur. Supposons que ϕ_{mod} soit une fonction de classe C^2 modélisant correctement la fonction photométrique ϕ décrite par (6.14). En suivant le même raisonnement que pour les images de luminance et en approximant la fonction ψ_{mod} par un développement en séries de Taylor, l'expression du modèle d'illumination de l'équation (6.31) s'exprime de la manière suivante :

$$\boxed{g(m') = \mathbf{f}(m) + \boldsymbol{\alpha}(x - x_p) + \boldsymbol{\beta}(y - y_p) + \boldsymbol{\gamma}} \quad (6.32)$$

Bien entendu, dans le cas d'un illuminant blanc, chacun des vecteurs $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ et $\boldsymbol{\gamma}$ possède des composantes égales.

Les modèles présentés ici sont adaptés très localement, notamment pour des zones pour lesquelles il n'y a pas de discontinuités très brutales de la courbure de surface ou de la rugosité. La principale limitation apparaît néanmoins lorsque des variations d'éclairage se produisent, puisqu'alors la réflectance ne doit pas montrer de fortes discontinuités. Or, ceci devient de moins en moins réaliste dès lors que la fenêtre d'intérêt est élargie. L'approximation de la réflectance par un polynôme de premier degré devient en effet de plus en plus grossière. Le prochain paragraphe répond précisément à ce problème.

6.3.2 Modèle d'illumination complet

D'après (6.20), λ dépend de l'angle d'incidence θ_i , qui peut varier fortement sur la fenêtre \mathcal{W} lorsque celle-ci est large ou lorsque la surface est non plane. De même, d'après (6.21), η dépend des variations de spéularité, des intensités et des valeurs d'angles d'incidence. En général, cette fonction n'est pas constante en tout point de \mathcal{W} .

Nous supposons que λ et η peuvent être modélisées sur \mathcal{W} par des fonctions de classe C^K , que nous appelons λ_{mod} et η_{mod} . Notons que contrairement à ϕ , la continuité λ et η ne suppose pas la continuité de $a(p)$. En outre, dans le cas où la surface est continue sur \mathcal{W} , cette hypothèse est vérifiée.

Images de luminance. Ainsi, nous développons les deux fonctions λ_{mod} et η_{mod} exprimées par (6.20) en séries de Taylor à l'ordre K autour de p :

$$\lambda_{mod}(m) = \sum_{i,j,i+j=0}^{i+j=K} \lambda_{ij}(p)(x-x_p)^i(y-y_p)^j \quad \eta_{mod}(m) = \sum_{i,j,i+j=0}^{i+j=K} \eta_{ij}(p)(x-x_p)^i(y-y_p)^j \quad (6.33)$$

Les termes λ_{ij} et η_{ij} correspondent aux dérivées partielles de λ_{mod} et de η_{mod} d'ordre i en x et j en y . En négligeant les termes des coefficients d'ordre élevé, ces fonctions sont approximées par :

$$\lambda_{mod}(m) \simeq \lambda_{mod}(p) + \left. \frac{\partial \lambda_{mod}}{\partial x} \right|_p (x - x_p) + \left. \frac{\partial \lambda_{mod}}{\partial y} \right|_p (y - y_p) \quad (6.34)$$

$$\eta_{mod}(m) \simeq \eta_{mod}(p) + \left. \frac{\partial \eta_{mod}}{\partial x} \right|_p (x - x_p) + \left. \frac{\partial \eta_{mod}}{\partial y} \right|_p (y - y_p) \quad (6.35)$$

Finalement, en notant $\lambda_1 = \left. \frac{\partial \lambda_{mod}}{\partial x} \right|_p$, $\lambda_2 = \left. \frac{\partial \lambda_{mod}}{\partial y} \right|_p$, $\lambda_3 = \lambda_{mod}(p)$, $\eta_1 = \left. \frac{\partial \eta_{mod}}{\partial x} \right|_p$, $\eta_2 = \left. \frac{\partial \eta_{mod}}{\partial y} \right|_p$ et $\eta_3 = \eta_{mod}(p)$, et en injectant (6.34) et (6.35) dans (6.19), la variation de luminance de m est donnée par :

$$\boxed{g(m') = [\lambda_1(x - x_p) + \lambda_2(y - y_p) + \lambda_3]f(m) + \eta_1(x - x_p) + \eta_2(y - y_p) + \eta_3} \quad (6.36)$$

Images couleur. Le modèle (6.36) se généralise également aux images couleur. Dans le cas d'un illuminant coloré, il devient :

$$\boxed{g(m') = [\lambda_1(x - x_p) + \lambda_2(y - y_p) + \lambda_3]f(m) + \eta_1(x - x_p) + \eta_2(y - y_p) + \eta_3} \quad (6.37)$$

Par contre, lorsque l'illuminant est blanc, la relation entre deux images f et g se simplifie puisque les paramètres du modèle possèdent des composantes égales et sont considérées comme des scalaires. Ce modèle peut prendre en compte différents types de changements d'illumination, que ce soit des variations de spéularité ou des variations d'éclairage.

Les modèles (6.31) et (6.36) (avec $\eta_1 = \eta_2 = \eta_3 = 0$) ont été utilisés dans [Black *et al.*, 2000] dans un contexte de correction d'image mais n'ont, à notre connaissance, pas été justifiés. Contrairement aux formalismes précédents, ils supposent que les changements d'illumination

peuvent être inégaux sur une même fenêtre d'intérêt \mathcal{W} . En particulier, le paramètre K_s et la rugosité n (ou ς) peuvent également varier. Par conséquent, les spécularités et les angles d'éclairage ne sont pas considérés constants sur \mathcal{W} .

Le nombre de paramètres à calculer dans (6.36) est plus important que pour les approches précédentes, mais le domaine de validité de cette modélisation s'étend à de larges zones \mathcal{W} de l'image. Contrairement à la modélisation proposée dans la section 6.3.1, celle-ci ne suppose pas la continuité de la réflectance du matériau (c'est-à-dire la continuité de $a(p)$), hypothèse difficilement admise pour de larges zones de l'image.

Dans la section qui suit, nous proposons d'analyser la validité de la modélisation (6.36) en fonction de la géométrie de la scène (position de la source et de la caméra) et des paramètres de la surface considérée (courbure et rugosité).

6.4 Validité des modèles locaux

Dans cette section, nous analysons la validité du modèle photométrique donné dans 6.3.2. Cette étude permettra également d'en déduire la validité du modèle proposé dans 6.3.1, dans le cas de variations de spécularité. Toutefois, afin de simplifier l'étude, nous la limitons au cas où la zone d'intérêt \mathcal{W} est proche de l'axe optique de la caméra.

6.4.1 Modélisation de la géométrie de la scène

Considérons un point P de l'objet de coordonnées $(X_p, Y_p, Z_p)^T$ dans un repère \mathcal{R}_c lié à la caméra, ce point est situé au centre d'une zone d'intérêt de l'objet. Soit également un point M situé dans le voisinage du point P (voir la figure 6.2). Dans le cas où la surface sur laquelle est positionné P est décrite localement par une fonction de classe C^2 , la profondeur de l'objet Z peut être développée en série de Taylor autour de P . Ses coordonnées (X, Y, Z) exprimées dans \mathcal{R}_c sont alors liées de la manière suivante :

$$Z = Z_p + D_X(X - X_p) + D_Y(Y - Y_p) + D_{XX}(X - X_p)^2 + D_{YY}(Y - Y_p)^2 + D_{XY}(X - X_p)(Y - Y_p) \quad (6.38)$$

où D_X, D_Y sont les dérivées premières de la surface au point P , décrivant ainsi l'orientation du plan tangent Π de la surface au point P par rapport au capteur :

$$D_X = \left. \frac{\partial Z}{\partial X} \right|_p \quad D_Y = \left. \frac{\partial Z}{\partial Y} \right|_p \quad (6.39)$$

Quant aux valeurs D_{XX}, D_{YY} et D_{XY} , elles correspondent aux dérivées secondes de la surface et témoignent de la courbure de la surface au point P :

$$D_{XX} = \left. \frac{\partial^2 Z}{\partial X^2} \right|_p \quad D_{YY} = \left. \frac{\partial^2 Z}{\partial Y^2} \right|_p \quad D_{XY} = \left. \frac{\partial^2 Z}{\partial X \partial Y} \right|_p \quad (6.40)$$

La normale à la surface au point M est alors déduite des dérivées au premier ordre de la surface au point M :

$$\mathbf{n} = \left(\frac{\partial Z}{\partial X}, \frac{\partial Z}{\partial Y}, -1 \right)^T \quad (6.41)$$

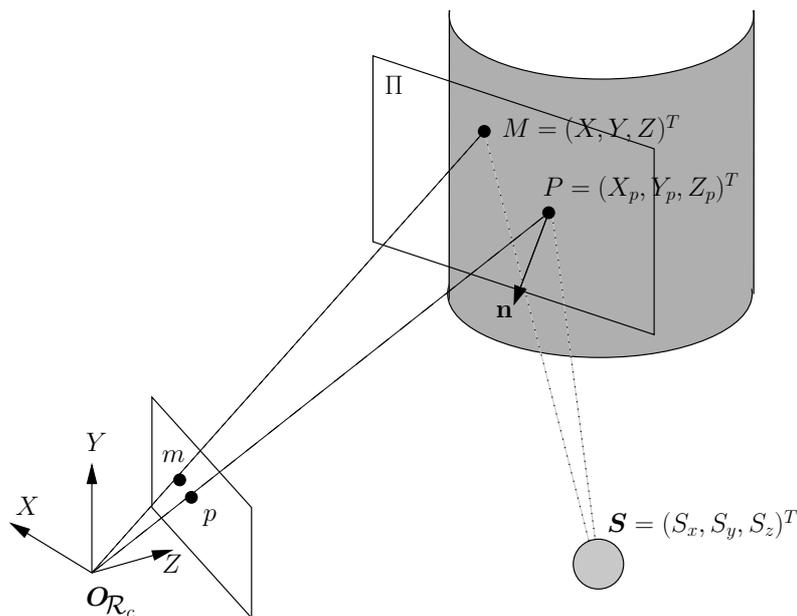


FIG. 6.2 – Modélisation de la géométrie de la scène.

En outre, on suppose que (6.38) est valide pour tout point de \mathcal{W} . Soit $\mathbf{S} = (S_x, S_y, S_z)^T$ le vecteur de position de l'éclairage ponctuel exprimé dans \mathcal{R}_c . On note $\mathbf{L} = (S_x - X, S_y - Y, S_z - Z)^T$ le vecteur directeur de l'éclairage dans ce même repère. Les termes de géométrie (l'angle θ_i notamment) peuvent alors être exprimés en fonction des coordonnées en pixels, après projection perspective.

Dans un premier temps, nous étudions la validité de l'approximation de la fonction λ (équation (6.20)), qui dépend de l'intensité et de l'angle d'incidence de l'éclairage. La modélisation proposée approxime la variation de cette fonction sur \mathcal{W} par l'équation d'un plan. Voyons dans quelles conditions cette approximation est justifiée, ou au contraire erronée.

6.4.2 Étude de la validité de la fonction λ

Pour ce faire, nous considérons la fonction λ décrite par (6.33), en se limitant à l'ordre $K = 2$:

$$\lambda(x, y) = \lambda_1 x + \lambda_2 y + \lambda_3 + \lambda_4 x^2 + \lambda_5 y^2 + \lambda_6 xy \quad (6.42)$$

Afin d'analyser la validité du modèle (6.36), nous étudions les conditions d'annulation des termes d'ordre 2 en x et y . Les configurations d'éclairage pour lesquelles ces termes sont négligeables témoignent des conditions pour lesquelles le modèle compense au mieux les changements d'illumination provoqués.

Pour ce faire, nous limitons l'étude au cas où la caméra et l'objet sont immobiles mais où un petit déplacement de l'éclairage direct $d\mathbf{S} = (dS_x, dS_y, dS_z)^T$ est provoqué par rapport à

sa position initiale \mathbf{S} , en considérant plusieurs positions d'observation et d'éclairage ainsi que différentes courbures de surfaces. En effet, le déplacement de la source d'éclairage permet de faire varier à la fois les deux termes λ et η . Les coefficients λ_i peuvent alors être développés en séries de Taylor autour de (dS_X, dS_Y, dS_Z) . Nous nous limitons au premier ordre en dS , de sorte de ne pas alourdir les expressions.

6.4.2.1 La direction d'éclairage coïncide avec la normale à la surface

Fixons le vecteur initial d'éclairage \mathbf{L} de sorte qu'il coïncide avec la normale \mathbf{n} à la surface au point P avec $\mathbf{L} = \alpha \mathbf{n}$. Pour de faibles variations de l'angle d'éclairage autour de la normale, les termes λ_4 , λ_5 et λ_6 sont nuls (leur développement en série de Taylor selon $d\mathbf{S}$ aboutit à des coefficients nuls). Par conséquent, l'approximation du changement d'illumination par notre modèle s'avère être très pertinente dans le cas où le vecteur d'éclairage est proche de la normale.

6.4.2.2 La source est proche de la caméra

Afin de traduire le cas de figure où la source d'éclairage et la caméra sont très proches, nous positionnons l'éclairage au niveau de l'origine du repère \mathcal{R}_c , tel que $\mathbf{S} = \mathbf{O}$. Le vecteur d'éclairage coïncide alors avec l'axe optique, ce qui n'est certes pas réaliste mais permet de formuler certaines conclusions. Nous considérons ici le cas d'un objet plan puis le cas d'un objet de forme non plane.

OBJET PLAN. Dans le cas d'un objet plan, les dérivées secondes (ou courbures) sont nulles, ce qui a l'avantage de simplifier grandement les équations. Les coefficients λ_i du second ordre deviennent :

$$\begin{cases} \lambda_{4(plan)} &= -\frac{2}{Z_P} dS_Z - \frac{2D_X}{Z_P} dS_X \\ \lambda_{5(plan)} &= -\frac{2D_Y}{Z_P} dS_Y - \frac{2}{Z_P} dS_Z \\ \lambda_{6(plan)} &= -\frac{D_X}{Z_P} dS_Y - \frac{D_Y}{Z_P} dS_X \end{cases} \quad (6.43)$$

Remarquons que, du fait de la présence de D_X et D_Y au numérateur, plus la surface est parallèle au plan du capteur ($D_X = D_Y = 0$) et moins les déplacements de la source pénalisent la validité du modèle. Cela est confirmé par les exemples de la figure 6.3, montrant les variations de λ (décrite par (6.20)) provoquées très localement dans une zone très restreinte \mathcal{W} de l'image. Une surface plane est considérée, et la source se déplace selon dS_X et dS_Y . Dans le cas où la surface de l'objet est parallèle au capteur, c'est-à-dire que $D_X = D_Y = 0$ (figure 6.3a) le modèle d'illumination est très bien adapté puisqu'il forme un plan dans la zone de l'image considérée \mathcal{W} . Par contre, lorsque $D_X \neq 0$ et $D_Y \neq 0$ (figure 6.3b) les variations d'illumination ne peuvent pas totalement être compensées par le modèle. Seuls des déplacements de la source dS_Z (un éloignement ou un rapprochement par rapport à l'objet) entraînent inévitablement des variations d'illumination qui ne sont pas compensées par notre modèle. Celles-ci sont d'autant plus importantes que la caméra (et par conséquent la source) est proche de la surface, comme le témoigne la présence de Z_P au dénominateur dans les relations (6.43).

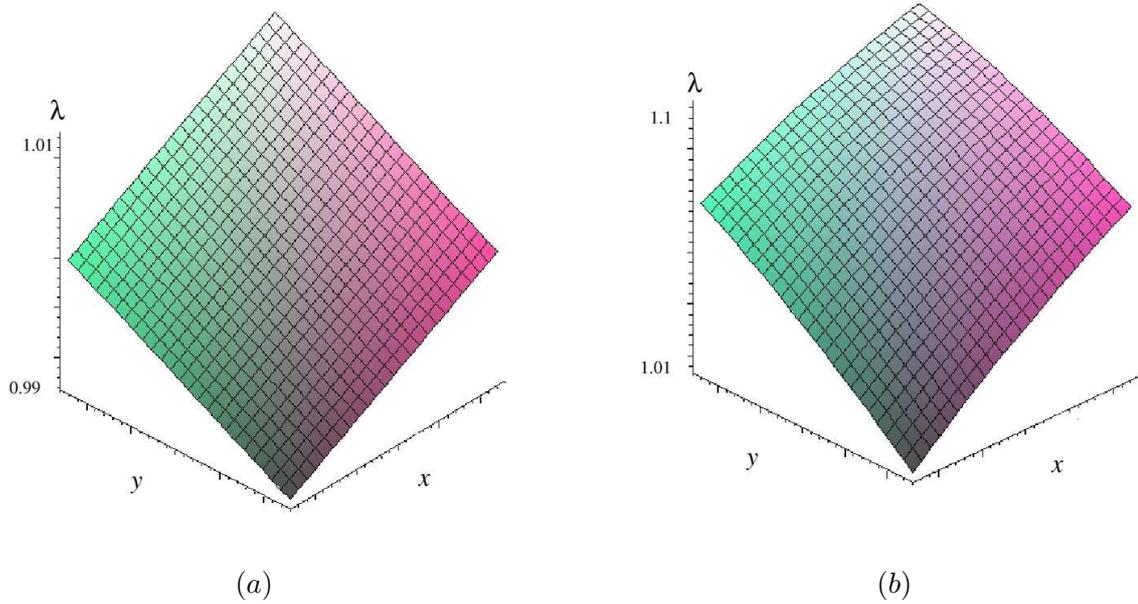


FIG. 6.3 – Exemples de variation d'illumination sur \mathcal{W} dans le cas où l'éclairage est proche de la caméra. (a) $D_X = D_Y = 0$, un déplacement de la source suivant X ou Y est compensé par le modèle d'illumination, qui forme effectivement un plan sur \mathcal{W} . (b) $D_X = D_Y = 5$, les variations d'illumination ne sont pas parfaitement compensées par le modèle. Dans les deux cas, les paramètres de modélisation sont les suivants : $D_{XX} = D_{YY} = D_{XY} = 0$, $X_P = Y_P = 0$, $Z_P = 100$ cm, $\mathbf{dS} = (0.1, 0.1, 0)^T$.

OBJET NON PLAN. Dans le cas d'un objet non plan pour lequel (6.38) est valide, les courbures de surfaces interviennent dans les paramètres de second ordre :

$$\begin{cases} \lambda_4 = \lambda_{4(plan)} + 2(D_Y D_{XX} + D_{XY} D_X) dS_Y - 4D_{XX} dS_Z + 6D_X D_{XX} dS_X \\ \lambda_5 = \lambda_{5(plan)} + 2(D_Y D_{XY} + D_X D_{YY}) dS_X + 6D_Y D_{YY} dS_Y - 4D_{YY} dS_Z \\ \lambda_6 = \lambda_{6(plan)} + 2(D_{YY} D_X + D_Y D_{XY}) dS_Y - 2D_{XY} dS_Z + 2(D_{XX} D_Y + D_X D_{XY}) dS_X \end{cases} \quad (6.44)$$

Plus les courbures de surfaces (D_{XX} , D_{XY} , D_{YY}) sont importantes et plus les coefficients λ sont susceptibles de varier en fonction d'un déplacement \mathbf{dS} de l'éclairage. Remarquons également qu'il suffit que l'orientation du plan tangent à la surface en P soit parallèle au capteur (D_X et D_Y nuls) pour que l'influence du déplacement d'éclairage dS_X et dS_Y soit annulée, quelle que soit la courbure de la surface. Par contre, un rapprochement de l'éclairage (ou inversement un éloignement) dS_Z a toujours une influence sur les termes du second degré. Par ailleurs, plus la surface est incurvée et plus cette influence est importante. Enfin, notons que la proximité de la caméra par rapport à la surface, décrite par Z_P , n'a pas plus d'influence sur les variations d'illumination lorsque la surface est non plane que dans le cas où la surface est plane.

6.4.2.3 La source est proche de la surface

Nous considérons à présent que la source d'éclairage est située à une petite distance ϵ de la surface. Après développement en séries de Taylor autour de $\epsilon = 0$, on obtient l'expression des coefficients λ_i suivante :

$$\left\{ \begin{array}{l} \lambda_4 = \frac{2Z_P(D_{XX}Z_P D_Y + D_{XY}D_X \epsilon)}{\epsilon^2} dS_Y - \frac{2Z_P(Z_P + D_X^2 \epsilon - Z_P D_X^2 + 2Z_P D_{XX} \epsilon)}{\epsilon^3} dS_Z + \\ \frac{2Z_P(D_X(\epsilon - 1) + D_{XX}(2\epsilon^2 + Z_P \epsilon))}{\epsilon^3} dS_X \\ \lambda_5 = \frac{2Z_P D_Y(D_{YY}Z_P \epsilon + 2D_{YY} \epsilon^2 - 2Z_P \epsilon)}{\epsilon^3} dS_Y - \frac{2Z_P(2D_{YY}Z_P \epsilon - D_Y^2 Z_P + Z_P + D_Y^2 \epsilon)}{\epsilon^3} dS_Z \\ + \frac{2Z_P(D_{YY}Z_P D_X + D_{XY}D_Y \epsilon)}{\epsilon^2} dS_X \\ \lambda_6 = \frac{Z_P(D_Y D_{XY}(Z_P \epsilon + \epsilon^2) + D_X(\epsilon + 2D_{YY} \epsilon^2 - 2Z_P))}{\epsilon^3} dS_Y - \frac{2Z_P(D_Y D_X(\epsilon - Z_P) + Z_P D_{XY} \epsilon)}{\epsilon^3} dS_Z + \\ \frac{Z_P(D_X D_{XY}(Z_P \epsilon + \epsilon^2) + D_Y(\epsilon + 2D_{XX} \epsilon^2 - 2Z_P))}{\epsilon^3} dS_X \end{array} \right. \quad (6.45)$$

Là encore, si la surface n'est pas orientée par rapport au capteur ($D_X = D_Y = 0$), un déplacement de l'éclairage dS_X et dS_Y n'apporte aucune contribution sur les valeurs λ_i . L'approximation des variations d'illumination par un plan est alors justifiée. En outre, elle s'avère la plus pertinente dans le cas où la source est plus éloignée de l'objet (ϵ grand) que ne l'est la caméra. Les contributions des variations dS_X, dS_Y, dS_Z aux termes λ_i sont alors minimales. En guise d'exemple, les figures 6.4a, 6.4b et 6.4c montrent les variations d'illumination provoquées par un déplacement de la source vis-à-vis de la surface. Dans le premier cas, l'éclairage est plus éloigné de la surface que ne l'est la caméra. Dans le second cas, la source et le capteur sont positionnés à une même distance, et enfin dans le troisième cas, la source est plus proche de la surface que ne l'est la caméra. Plus la source se rapproche de la surface et moins le modèle proposé est pertinent.

6.4.2.4 Conclusion

En définitive, les configurations idéales de validité de l'approximation de λ par un développement de Taylor au premier ordre sont obtenues dans les cas suivants :

- le plan tangent Π à la surface en P est peu incliné par rapport au plan du capteur, et les courbures de la surface de l'objet sont faibles. Ces deux points illustrent surtout le fait que la surface ne doit pas montrer de discontinuités de profondeur trop fortes ;
- la direction d'éclairage L est concourante à la normale \mathbf{n} au point considéré ;
- la source est plus éloignée de la surface que ne l'est le capteur.

Toutefois, cette approximation s'avère plus adéquate qu'une approximation par une constante qui requiert l'annulation des termes de second ordre et du premier ordre. Or, comme on pu le montrer les quelques exemples de variations d'illumination fournis, elles n'apparaissent en aucun cas constantes.

6.4.3 Étude de la validité de la fonction η

De manière à étudier le domaine de validité de la fonction η , exprimée par (6.20), il est nécessaire de prendre en compte la modélisation des spécularités, et par conséquent les propriétés

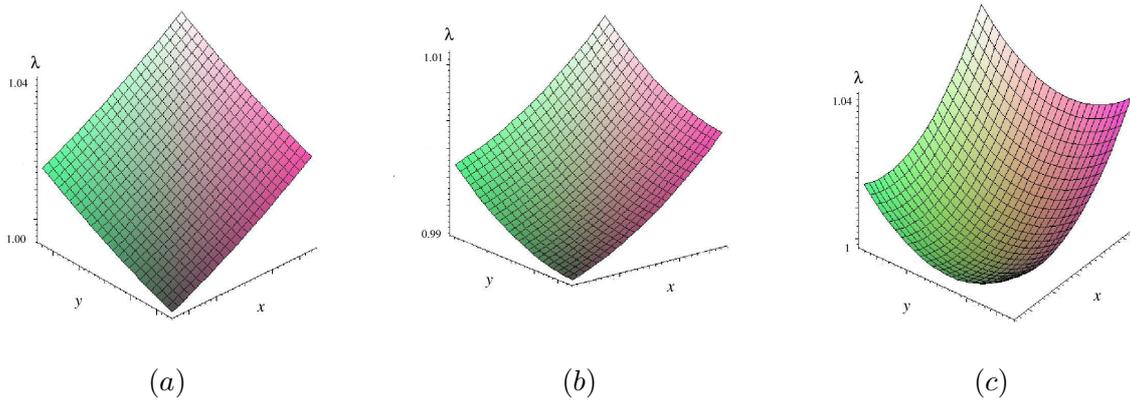


FIG. 6.4 – Exemples de variation d’illumination provoquée sur \mathcal{W} dans le cas où l’éclairage est proche de la surface. (a) La source est plus éloignée de la surface que ne l’est la caméra ($S_Z = 15$, $Z_P = 10$ cm). (b) La source et la caméra se situent à une même distance de la surface ($S_Z = Z_P = 10$). (c) La source est plus proche de la surface que ne l’est la caméra ($S_Z = 2$, $Z_P = 10$). Dans les trois cas, les paramètres utilisés sont les suivants : $\alpha = 0, 2$, $X_P = Y_P = 0$, $D_Y = D_{XY} = 0$, $D_X = D_{XX} = D_{YY} = 0.1$, $d\mathbf{S} = (1, 1, -0.5)^T$.

de l’objet. Pour ce faire, nous considérons le modèle spéculaire de Phong (6.5). Afin de simplifier les équations, nous considérons que le terme $\lambda(m)$ est égal à l’unité et qu’il n’y a donc pas de variation d’intensité d’éclairage ni de changement de la géométrie de la scène. Dans ce cas, η équivaut à la fonction ψ décrite par (6.10). Nous étudions ainsi la validité de l’expression suivante :

$$\eta(m') = h_g(M) - h_f(M) \quad (6.46)$$

Après développement en série de Taylor au second ordre, η est approximée par :

$$\eta(m) = \eta_1 x + \eta_2 y + \eta_3 + \eta_4 x^2 + \eta_5 y^2 + \eta_6 xy \quad (6.47)$$

où les termes η_i dépendent des termes de géométrie introduits dans la section 6.4.1. Puisque les fonctions de spécularité h montrent un maximum lorsque l’angle ρ entre la bissectrice \mathbf{B} entre \mathbf{L} et \mathbf{V} est nul, il s’avère intéressant d’étudier la validité des modélisations dans ce cas précis. La position initiale de la source est choisie de sorte que la normale à la surface \mathbf{n} soit concourante à \mathbf{B} (voir la figure 6.5). On note α la distance entre la source et le point P . Nous admettons là encore un petit déplacement de l’éclairage $d\mathbf{S}$, ce qui permet d’effectuer un développement limité des coefficients η_i autour de ces déplacements. Nous étudions ici quelques cas particuliers de configuration de la scène.

6.4.3.1 Vecteurs d’éclairage, d’observation et vecteur normal concourants

À l’instant initial (avant déplacement de la source), les vecteurs \mathbf{L} , \mathbf{V} et \mathbf{n} sont considérés concourants. Ainsi, le plan tangent à la surface est parallèle au plan du capteur et $D_X = D_Y = 0$ et l’angle d’éclairage θ_i est nul. Nous considérons ici un objet non plan donné par (6.38). L’expression des coefficients η_4 , η_5 et η_6 étant trop complexe pour pouvoir être analysée, nous étudions quelques configurations particulières.

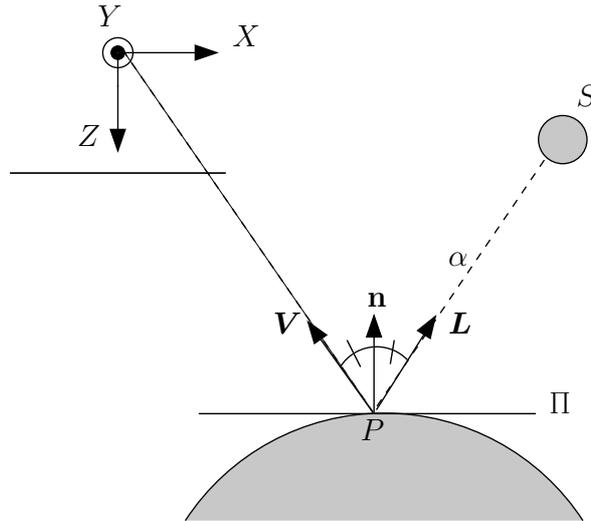


FIG. 6.5 – La normale à la surface en P correspond à la bissectrice entre s et v .

1-*L'éclairage est proche de la surface.* Dans le cas où l'éclairage est proche de la surface, c'est-à-dire lorsque α est faible, les coefficients η_i peuvent être développés en séries de Taylor autour de $\alpha = 0$. Après calculs, au premier ordre, les valeurs η_i sont simplifiées comme suit :

$$\begin{cases} \eta_4 = -n \left(2D_{XX} + \frac{1}{Z_P} \right) dSZ \\ \eta_5 = -n \left(2D_{YY} + \frac{1}{Z_P} \right) dSZ \\ \eta_6 = -n D_{XY} dSZ \end{cases} \quad (6.48)$$

Dans le cas où l'éclairage est proche de la surface, un rapprochement (ou un éloignement) dS_Z de la source par rapport à la surface entraîne toujours des variations des η_i , que la surface soit plane ou non. Par contre, le paramètre η_6 est nul dès lors que $D_{XY} = 0$. C'est notamment le cas des surfaces planes ou des surfaces de révolution. Le déplacement selon Z aura moins d'influence si le capteur est suffisamment éloigné de la surface, et si la surface est rugueuse (n faible) et plane ($D_{XX} = D_{YY} = D_{XY} = 0$).

2-*La caméra est proche de la surface (Z_P faible).* Dans ce cas on peut montrer que tous les coefficients sont nuls et la modélisation des changements de réflexion est justifiée. La figure 6.6 illustre ce cas de figure, où la surface considérée est non plane. Les variations de η sont bien compensées par un développement de Taylor au premier ordre.

6.4.3.2 Petite inclinaison de la surface par rapport au plan du capteur

Dans le cas d'une petite inclinaison de la surface par rapport au capteur (D_X et D_Y faibles), les coefficients η_i peuvent être développés en séries de Taylor autour de $D_X = D_Y = 0$. On considère toujours que la normale à la surface n coïncide avec la bissectrice B entre L et V de sorte que plus l'orientation augmente et plus l'angle d'incidence θ_i augmente. Nous nous

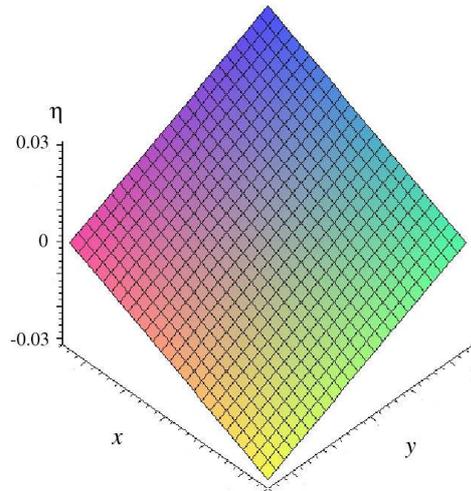


FIG. 6.6 – Exemple de variation de η lorsque la caméra est proche de la surface $Z_P = 10$. L'objet est non plan, $D_X=0$, $D_Y=0$, $D_{XX}=0.1$, $D_{YY}=0.1$, $D_{XY}=0$. Le déplacement de la source vaut $d\mathbf{S} = (1, -1, -1)^T$ et $\alpha = 100$.

intéressons ici exclusivement au cas des objets plans, étant donnée la complexité des expressions dans le cas des objets non plans. Là encore, différentes configurations d'éclairage sont considérées.

1-L'éclairage est proche de la surface. Les expressions de η_4 , η_5 et η_6 sont données par :

$$\begin{cases} \eta_4 &= -\frac{nD_X}{4Z_P} (3n + 7) dS_X - \frac{nD_X}{4Z_P} (n + 1) dS_Y - \frac{n}{Z_P} dS_Z \\ \eta_5 &= -\frac{nD_Y}{4Z_P} (n + 1) dS_X - \frac{nD_Y}{4Z_P} (3n + 7) dS_Y - \frac{n}{Z_P} dS_Z \\ \eta_6 &= \frac{nD_X(n+3)}{4Z_P} dS_X + \frac{nD_Y(n+3)}{4Z_P} dS_Y \end{cases} \quad (6.49)$$

Dès lors que D_X et D_Y sont non nuls, et donc θ_i également, un déplacement de la source (dS_X , dS_Y) provoque des variations des η_i . Celles-ci augmentent d'autant plus que le matériau est lisse (n élevé), que la caméra est proche (Z_P faible) et que l'inclinaison du plan tangent par rapport au capteur est significative (D_X et D_Y importants). Ce dernier point est illustré par les figures 6.7a et 6.7b, qui montrent respectivement un exemple de variation de η lorsque le plan tangent de la surface au point P est faiblement (figure 6.7a) et fortement orienté (figure 6.7b).

2-La caméra est proche de la surface. Dans le cas où la surface est plane, légèrement inclinée par rapport au plan du capteur, et que la caméra est suffisamment proche de la surface (Z_P faible), alors l'approximation de η par un plan est parfaitement justifiée. En effet, les termes η_4 , η_5 et η_6 sont tous négligeables. C'est ce qu'illustre la figure 6.8, qui montre un exemple de variation de η sur une zone de l'image \mathcal{W} .

6.4.3.3 Conclusion

Dans le cas où $\lambda = 1$, l'approximation du terme η (qui est alors équivalent à ψ) par un polynôme du premier degré se justifie au mieux lorsque les conditions suivantes sont respectées :

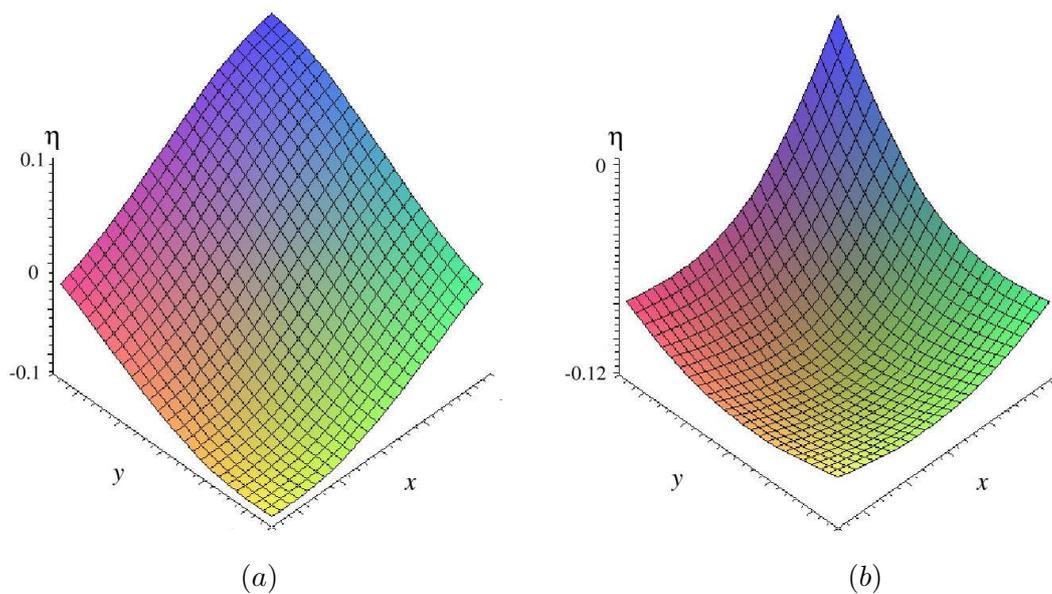


FIG. 6.7 – Exemples de variation de η et que la source est proche de la surface, lorsque le plan tangent à la surface au point P est peu incliné (figure (a)) ou très incliné (figure (b)) par rapport au plan capteur.

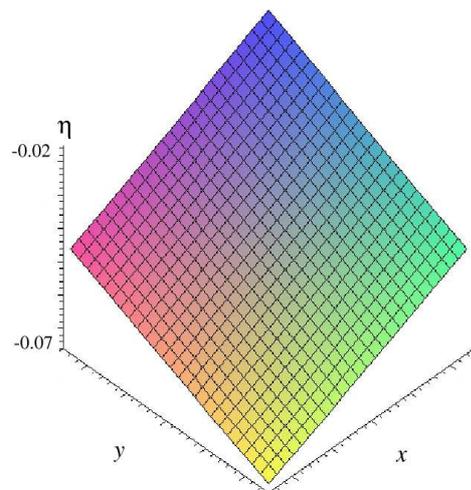


FIG. 6.8 – Exemple de variation de η lorsque le plan tangent de l'objet n'est pas incliné par rapport au plan capteur, et que la caméra est proche de la surface.

- la surface considérée est peu incurvée et le plan tangent au point considéré est peu incliné par rapport au plan du capteur ;
- la surface est rugueuse ;
- la caméra est proche de la surface.

Les variations de spécularité, qui forment approximativement un plan sur \mathcal{W} sont alors correctement compensées. En prenant en compte les remarques de la section 6.4.2.4 page 140, notons

en définitive que les approximations de λ et η sont d'autant plus précises que la source est éloignée de la surface et que la caméra est proche de la surface. D'autre part, il est préférable que la surface ne soit pas discontinue ou excessivement lisse.

En revanche, les modèles photométriques décrits dans la section 6.2 reposent sur des hypothèses plus restrictives, où les fonctions λ et η sont considérées constantes en tout point de \mathcal{W} . Non seulement les coefficients de Taylor au second ordre dans (6.42) et (6.47) sont nuls, mais également une partie des coefficients du premier ordre ($\lambda_2 = \lambda_3 = \eta_2 = \eta_3 = 0$). Cependant, ces termes dépendent de la courbure de la surface considérée, de la géométrie de la scène (positionnement relatif entre la caméra, la surface et la source d'éclairage) et sont donc susceptibles de varier sur \mathcal{W} . Quelques exemples de variations d'illumination (de la figure 6.3a à la figure 6.8) confirment ces propos.

6.5 Conclusion

À partir des modèles de réflexion largement utilisés en vision par ordinateur [Phong, 1975, Torrance et Sparrow, 1967], nous avons exprimé clairement les variations d'illumination susceptibles d'apparaître entre deux images. Celles-ci dépendent de nombreux facteurs, à savoir les variations d'intensité et/ou de couleur de l'illuminant (de l'éclairage ambiant ou de l'éclairage ponctuel), les changements d'angle d'incidence de la lumière, d'angle d'observation, la rugosité du matériau et la courbure de la surface observée. Cette modélisation photométrique met en jeu un si grand nombre de paramètres que son utilisation s'avère difficile dans une analyse de séquence d'images ou dans une optique de mise en correspondance, à moins de disposer de connaissances *a priori* sur le contenu de la scène. Par contre, des approximations locales, telles que les modèles additif, scalaire et affine, sont plus aisément exploitables [Gros, 2000, Jin *et al.*, 2001]. Ce chapitre a notamment permis d'expliquer sur quelles hypothèses précises se basent ces modèles. Toutefois, pour la plupart d'entre eux, les changements d'illumination sont considérés localement constants, ce qui n'est en général pas le cas, surtout lorsque la surface considérée n'est pas plane.

Deux modèles présentés ici permettent néanmoins de compenser les variations spatiales de changements d'illumination. Le premier formalisme compense de manière très locale les variations photométriques et sous-entend, dans le cas de changements d'éclairage, la modélisation de la réflectance par une fonction continue. Le second s'avère plus complet et ne modélise que les variations spatiales dues à la géométrie. Il s'avère donc mieux approprié à une modélisation plus globale des changements d'illumination. Là encore, leur fondement théorique a été justifié en se basant sur les modèles de réflexion couramment utilisés en vision par ordinateur.

Enfin, une partie de ce chapitre a concerné l'étude de la validité de ces modèles en fonction de la configuration de la scène et des propriétés de la surface. La compensation des variations de spécularité par le modèle local décrit dans la section 6.3.1 a par la même occasion été analysée.

Dans le chapitre qui suit, nous nous intéressons à l'utilisation de ces modèles dans le cadre du suivi dans les images de luminance.

SUIVI DE PRIMITIVES ROBUSTE AUX VARIATIONS D'ILLUMINATION DANS LES IMAGES DE LUMINANCE

Ce chapitre aborde la problématique du suivi de points et de motifs plus larges vis-à-vis des variations d'illumination. Il se réfère largement au chapitre 6 quant à l'utilisation des modèles de variation d'illumination, dont l'intégration dans le schéma de suivi permet d'assurer une certaine robustesse. Après un état de l'art des approches de suivi de points, nous proposons d'améliorer le suivi par la prise en compte des modèles photométriques décrits dans la section 6.2. Nous comparons cette approche avec les techniques existantes par le biais d'expérimentations effectuées sur des objets spéculaires, et pour des séquences montrant des variations d'éclairage.



De nombreuses applications de vision par ordinateur requièrent l'appariement ou la mise en correspondance de primitives entre deux prises de vue. C'est le cas des méthodes de calcul de mouvement 2D, du suivi d'objet ou de cible, la vision dynamique, active, ou encore l'asservissement visuel. Dans le cadre applicatif agro-alimentaire dans lequel s'inscrivent ces travaux de thèse, il s'agit d'effectuer un positionnement de la caméra par rapport à l'objet, par asservissement visuel, de manière à en assurer l'inspection, ou de lire un identifiant (voir le chapitre 5). Or, les scènes considérées sont de type peu structuré [Alhaj, 2004]. Dans ce contexte, les primitives de type contours sont difficilement exploitables. Comme nous l'avons souligné dans l'introduction de cette partie (page 123), des primitives telles que les points et les régions texturées de l'image sont les plus facilement extractibles. Puisqu'elles dépendent directement des luminances de l'image, et pas de la structure des objets dans la scène (comme les contours par exemple), le suivi de telles primitives peut souffrir des variations d'illumination. Dans ce chapitre, nous proposons d'améliorer la robustesse du suivi par la prise en compte des variations d'illumination, en se basant sur les modèles photométriques proposés dans le chapitre 6.

La section 7.1 aborde les principales techniques de détection de points d'intérêt existantes. Ensuite, la section 7.2 dresse un état de l'art concernant les techniques de suivi de points d'intérêt, décrivant deux types de techniques : les approches *corrélatives* et les approches *différentielles*. Nous consacrons la section 7.3 aux méthodes de suivi différentielles compensant en partie les variations d'illumination. La méthode proposée, basée sur l'utilisation d'un modèle d'illumination prenant en compte les variations spatiales de changement d'illumination est décrite dans la section 7.4. Elle est étendue au cas de motifs plus larges dans la section 7.5. Enfin, les approches proposées et les techniques existantes sont comparées dans la section 7.6, à l'aide de séquences simulées et réelles où apparaissent des spécularités et/ou des variations d'éclairage. Notons que les notations prises en compte dans le chapitre 6 sont conservées.

7.1 Détection de points d'intérêt

Les approches de détection de points peuvent être classées suivant le type d'images sur lesquelles elles s'appliquent, selon qu'elles sont structurées ou non. Le lecteur pourra trouver dans [Schmid *et al.*, 2000] un état de l'art plus complet de ces approches.

7.1.1 Méthodes de détection applicables aux images structurées

Les images structurées, telles que celles montrant des scènes d'intérieur ou d'objets manufacturiers, présentent généralement des contours bien marqués entre les différentes régions de l'image, formant des coins ou encore des jonctions en forme de X, Y ou T. C'est en se basant sur ce constat que certains auteurs ont proposé de détecter des points caractéristiques après une extraction préalable des contours. Les changements de courbure de ces lignes [Asada et Brady, 1986], leurs points d'inflexion [Mokhtarian et Mackworth, 1986], les *maxima* de courbures [Medioni et Yasumoto, 1987] ou encore l'intersection de plusieurs contours [Horaud *et al.*, 1990] déterminent alors la position des points d'intérêt. Or, la détection de frontières est souvent imprécise dans les zones d'inflexion de ces contours, ce qui nuit finalement à la précision de la détection des coins. C'est pourquoi dans [Chen *et al.*, 1995], cette tâche est améliorée en exploitant les propriétés de la transformation en ondelettes. Notons également les travaux de [Yeh, 2003], basés également sur le calcul d'ondelettes.

D'autres approches définissent un modèle théorique du signal caractérisant le point d'intérêt, ce qui autorise la recherche et la reconnaissance du point dans l'image. Il existe ainsi des modèles théoriques de coins, de jonction de contours en forme de T. Dans [Rohr, 2002] par exemple, le point est défini par un modèle de jonction, paramétré notamment par ses niveaux de gris, l'ouverture du coin, ainsi que le niveau de flou. Une méthode de minimisation de type Levenberg-Marquardt permet alors d'identifier le point, ainsi que ses paramètres.

Une tendance plus récente consiste à utiliser les outils de la morphologie mathématique pour extraire et reconnaître différentes formes de coins [Laganiere, 1998], qui sont généralement exploités dans un cadre de reconnaissance de formes. Ce type d'approche a récemment été amélioré pour détecter des angles très obtus, et également réduire le temps de calcul [Shih *et al.*, 2005].

Ces différentes approches de détection de points d'intérêt nécessitent une certaine structure de l'image, par la présence de contours significatifs, de coins ou de formes géométriques très marquées. De ce fait, ces primitives sont peu dépendantes des changements d'illumination et de la géométrie de la scène. Or, un grand nombre d'images naturelles, dépourvues de tels agencements de pixels, ne peuvent être caractérisées que par leur luminance.

7.1.2 Méthodes de détection applicables aux images texturées

Afin que des points caractéristiques soient détectés dans des images peu structurées, il est nécessaire de pouvoir les extraire directement à partir du signal. Depuis la fin des années 1970, un grand nombre de détecteurs a été proposé, tel que dans [Beaudet, 1978, Kitchen et Rosenfeld, 1982] ou encore la très populaire technique de Moravec [Moravec, 1979]. Cette dernière se base sur une fonction d'autocorrélation du signal. Elle mesure les différences de luminance entre une fenêtre d'intérêt et des fenêtres déplacées dans quatre directions discrètes de l'image (suivant x et y). Les fenêtres d'intensité homogène obtiennent une valeur d'autocorrélation faible dans toutes les directions, tandis que les zones de frontières unidirectionnelles mènent à une valeur élevée dans une seule des directions. Enfin, les primitives à détecter ont une valeur d'autocorrélation élevée quelle que soit la direction.

Cette technique est améliorée plus tard par Harris [Harris et Stephens, 1988] et par Kanade, Lucas et Tomasi [Lucas et Kanade, 1981, Tomasi et Kanade, 1991], méthode couramment dénommée KLT. Tout d'abord, f est approximée en série de Taylor au voisinage du point p , afin de remédier au caractère anisotropique des directions prises en compte par la technique de Moravec. De cette manière, la fonction d'autocorrélation est simplifiée par le calcul d'une matrice M en chaque point p de l'image :

$$M(p) = \sum_{m \in \mathcal{W}} \begin{pmatrix} f_x^2(m) & f_x(m)f_y(m) \\ f_x(m)f_y(m) & f_y^2(m) \end{pmatrix} \quad (7.1)$$

Les termes f_x et f_y correspondent aux dérivées premières de f selon les directions x et y . Ainsi M caractérise la structure locale autour du point p . Les valeurs propres de M , ν_1 et ν_2 , témoignent du profil d'intensité de la fenêtre \mathcal{W} . Le détecteur de Harris et le détecteur KLT se distinguent par la détection des points caractéristiques à partir des valeurs de M .

Le détecteur de Harris calcule un critère $M_f(p)$ basé sur le déterminant « det » et la trace « tr » de $M(p)$:

$$M_f(p) = \det M(p) - \kappa(\text{tr} M(p))^2 \quad (7.2)$$

où κ est un paramètre fixé par l'utilisateur. La valeur $M_f(p)$ favorise la sélection des points pour lesquels $M(p)$ possède deux valeurs propres élevées et d'importance comparable. Dans le cas d'un point caractéristique, $M_f(p)$ est supérieure à 0.

Le détecteur KLT a été développé dans une optique de suivi de points. Les points détectés sont les points susceptibles d'être suivis, c'est-à-dire pour lesquels la matrice M est bien conditionnée, c'est-à-dire si $\max(\nu_1, \nu_2) < \nu_s \min(\nu_1, \nu_2)$.

Dans [Trajkovic et Hedley, 1998], les auteurs proposent une version plus rapide du détecteur de Harris. En comparant l'approche de Harris avec d'autres techniques, ils attestent également

de sa bonne stabilité et de sa bonne précision. D'après d'autres travaux ayant comparé différents détecteurs de points dans des contextes de mise en correspondance [Schmid *et al.*, 2000, Vincent et Laganière, 2001] ou de suivi [Tissainayagam et Suter, 2004], le détecteur de Harris s'avère très performant d'un point de vue de la stabilité, de la répétabilité, de la robustesse au bruit. Notons que dans [Faille, 2004], l'auteur améliore la répétabilité de cette approche vis-à-vis des variations d'illumination.

Dans les images peu structurées, les primitives de type point sont les plus facilement extractibles. En revanche, le suivi ou plus précisément la détermination de la trajectoire d'un point dans l'image s'avère une tâche délicate. En effet, la seule information disponible devient la luminance du point observé ainsi que celle de quelques autres points situés dans son proche voisinage.

7.2 Suivi de points d'intérêt

Les techniques de suivi de points d'intérêt peuvent être regroupées en deux classes, selon qu'elles requièrent une recherche exhaustive par calcul d'une fonction de similarité (ou de dissimilarité) ou la minimisation d'une fonction de coût. On distingue ainsi les approches dites de *corrélacion* des approches dites *différentielles*.

7.2.1 Méthodes de corrélacion

Les méthodes de suivi par corrélacion considèrent un voisinage \mathcal{W} de taille $\mathcal{N} \times \mathcal{N}$ d'un point p dans une image f et recherchent dans une seconde image g la position du point p' dont le voisinage \mathcal{W} permet de maximiser une fonction de similarité (mesures de corrélacion) ou minimiser une fonction de dissimilarité (emploi de distances). Les fonctions de dissimilarité les plus utilisées sont la somme des différences absolues (SAD), la somme des distances au carré (SSD); la mesure de similarité la plus courante est la corrélacion croisée (CC), existant également dans une version normalisée (NCC). L'ensemble de ces critères de corrélacion peuvent être centrés, menant aux critères ZSAD, ZSSD, ZNCC. Ainsi, le critère ZNCC s'écrit :

$$ZNCC(p, p') = \frac{1}{N} \sum_{p+dp, p'+dp \in \mathcal{W}} \frac{[f(p+dp) - \mu_f(p)] [g(p'+dp) - \mu_g(p')]}{\sigma_f \sigma_g} \quad (7.3)$$

où dp correspond à un incrément en pixels (d_x, d_y) par rapport aux positions p et p' . Les valeurs μ_f et μ_g correspondent aux moyennes des intensités appartenant aux fenêtres centrées respectivement sur p et p' dans les images f et g . Les termes σ_f et σ_g expriment quant à eux les écarts-type. Nous montrerons dans la suite du document, tout l'intérêt de cette normalisation photométrique. Ce type de mesures corrélatives est également largement utilisé en stéréovision [Banks et Corke, 2001].

La recherche dans l'image g de la fenêtre d'intérêt la plus similaire à la fenêtre d'intérêt initiale peut s'accomplir de manière exhaustive dans toute l'image g , ou bien dans un espace de recherche plus restreint, centré autour de p , tel que l'illustre la figure 7.1. Le choix de la taille de la zone de recherche s'avère crucial au bon fonctionnement de la méthode de suivi.

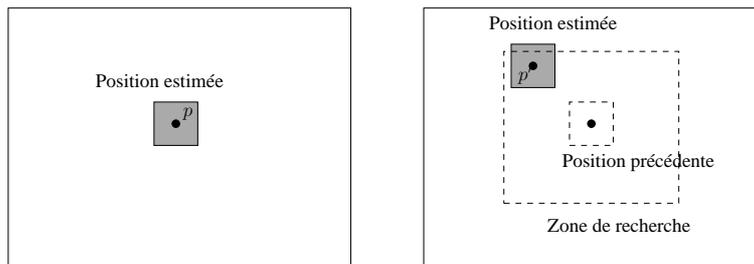


FIG. 7.1 – Illustration de la méthode de corrélation.

Elle doit être suffisamment large pour ne pas perdre les points lorsque les déplacements inter-images sont importants, mais doit être suffisamment restreinte de manière à éviter les ambiguïtés éventuelles. Pour finir, elle ne doit pas être trop coûteuse en temps de calcul.

Une des limitations des méthodes de corrélation réside dans leur faible prise en compte des déformations géométriques imputables au mouvement de la caméra ou de l'objet. Généralement, de simples translations ou similitudes sont considérées. Néanmoins, les changements d'échelle peuvent être compensés par modélisation [Seelen et Bajcsy, 1997] ou par une variation de la taille de fenêtre, tel que dans [Kanade et Okutomi, 1994, Chien et Sung, 2000], ce type d'approches permettant en outre de prendre en compte les variations de forme de l'objet à suivre. Afin de contourner le problème de déformation, [Matthews *et al.*, 2004] propose une méthode simple de remise à jour de la primitive à suivre, en évitant les dérives temporelles. La primitive n'est mise à jour que si les erreurs de corrélation sont suffisamment faibles par rapport à la primitive initiale. Notons également qu'ils utilisent une déformation (similitude de l'image) et une modélisation des variations d'illumination [Black *et al.*, 2000].

Contrairement aux approches de suivi de points par corrélation, les méthodes différentielles, que nous décrivons dans la section suivante, permettent de compenser des déformations géométriques plus significatives.

7.2.2 Méthodes différentielles

Contrairement aux méthodes de suivi par corrélation, les méthodes différentielles ne requièrent pas de recherche exhaustive. Le mouvement de la fenêtre d'intérêt \mathcal{W} centrée autour d'un point p entre deux images est modélisée, et les paramètres de ce modèle sont calculés par minimisation d'une fonction de coût. Les travaux précurseurs en ce domaine sont dus à Lucas et Kanade [Lucas et Kanade, 1981, Tomasi et Kanade, 1991]. Leur approche repose sur la contrainte de flot optique exprimée par [Horn et Schunck, 1981], c'est-à-dire sur l'hypothèse de conservation de la luminance du point observé au cours de son déplacement (voir le paragraphe 6.2.1 page 131). Encore aujourd'hui, cette technique KLT, que nous allons décrire dans la section 7.2.2.1, est réputée très performante [Tissainayagam et Suter, 2004], et il en existe de nombreuses variantes, dont une comparaison est effectuée dans [Baker, 2004].

7.2.2.1 Méthode classique

Soit f et g deux images d'une même scène acquises selon deux points de vue différents. Supposons que la déformation de f attribuée au mouvement de la caméra et/ou de l'objet soit représentée en tout point m par un mouvement dans l'image décrit par la fonction δ , telle que $m' = \delta(m, \mathbf{A})$:

$$f(m) = g(\delta(m, \mathbf{A})) \quad (7.4)$$

L'algorithme de suivi consiste à calculer les paramètres \mathbf{A} du modèle de déplacement que subit le point m entre les deux images f et g . Or, l'équation (7.4) n'est pas exacte à cause de l'apparition de bruit, du changement éventuel d'illumination mais également parce que le modèle de déplacement n'est qu'une approximation. C'est pourquoi il est préférable de calculer le mouvement entre l'image à l'instant k et celle à l'instant k' par minimisation de la fonction de coût suivante :

$$\epsilon_1(\mathbf{A}) = \sum_{m \in \mathcal{W}} (f(m) - g(\delta(m, \mathbf{A})))^2 \quad (7.5)$$

Pour obtenir \mathbf{A} , on suppose que $\mathbf{A} = \hat{\mathbf{A}} + \mathbf{a}$ où \mathbf{a} exprime une petite variation autour d'une estimation $\hat{\mathbf{A}}$ de \mathbf{A} [Collewet *et al.*, 2004]. Dans ces conditions, $g(\delta(m, \mathbf{A}))$ peut être développé en série de Taylor au 1^{er} ordre autour de $\hat{\mathbf{A}}$:

$$g(\delta(m, \mathbf{A})) = g(\delta(m, \hat{\mathbf{A}})) + \nabla g^T(\delta(m, \hat{\mathbf{A}})) J_\delta^{\mathbf{A}} \mathbf{a} \quad (7.6)$$

où $J_\delta^{\mathbf{A}}$ représente la matrice Jacobienne de la fonction δ par rapport à \mathbf{A} . En injectant ensuite (7.6) dans (7.5), on aboutit finalement à un système linéaire en \mathbf{a} . Celui-ci se résout selon un schéma itératif afin de prendre en compte les approximations introduites par le développement en série de Taylor. Tous calculs faits, \mathbf{a} est obtenu par résolution du système linéaire suivant :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V} \mathbf{V}^T \right) \mathbf{a} = \sum_{m \in \mathcal{W}} \left(f(m) - g(\delta(m, \hat{\mathbf{A}})) \right) \mathbf{V} \quad (7.7)$$

où \mathbf{V} est le vecteur donné par :

$$\mathbf{V} = (J_\delta^{\mathbf{A}})^T \nabla g(\delta(m, \hat{\mathbf{A}})). \quad (7.8)$$

Décrivons à présent le modèle de mouvement utilisé dans le suivi de points.

7.2.2.2 Modélisation des déformations géométriques

Les travaux précurseurs de [Lucas et Kanade, 1981] emploient un modèle de mouvement translationnel, défini en tout point m de \mathcal{W} de la manière qui suit :

$$m' = m + \mathbf{T} \quad (7.9)$$

Il suppose que tout point m de la fenêtre d'intérêt \mathcal{W} subit une translation dont le vecteur est identique en chaque point de \mathcal{W} . Or ce cas de figure peut se produire dans des cas très particuliers où la surface observée se déplace de manière fronto-parallèle. Les rotations ne sont pas

prises en compte. Du fait des hypothèses restrictives sur lesquelles repose ce modèle de déformation, ses paramètres ne peuvent être approximés que lorsque le mouvement entre les deux images f et g est suffisamment faible. Deux images successives sont généralement considérées. Or, cela n'est pas sans conséquences, puisque cela peut entraîner une accumulation d'erreurs de suivi tout au long de la séquence, et donc des dérives du point suivi. La robustesse du suivi a pu être améliorée en considérant un modèle affine [Shi et Tomasi, 1994] :

$$m' = \mathbf{D}m + \mathbf{T} \quad (7.10)$$

Du fait de la bonne adéquation de ce modèle entre robustesse, simplicité et précision, il est très largement utilisé en suivi de points. Il permet de modéliser de nombreuses déformations géométriques, telles que des rotations, des homothéties, des translations. Contrairement au modèle translationnel, le modèle affine s'avère suffisamment réaliste pour que ses paramètres puissent être calculés entre l'image initiale de la séquence f et l'image courante g . Les problèmes de dérives sont alors écartés. Ce modèle sera utilisé dans la suite de ce document.

Il existe d'autres modèles de mouvement plus complets, tels que le modèle quadratique [Odobez et Bouthemy, 1995] et le modèle homographique [Buenaposada et Baumela, 2002, Collewet *et al.*, 2004]. Du fait du grand nombre de paramètres mis en jeu, ces modèles sont peu exploités dans les applications de suivi de points d'intérêt, où les fenêtres d'intérêt \mathcal{W} sont de taille restreinte.

Améliorer la précision du modèle de mouvement n'est pas la seule façon d'assurer la robustesse du suivi. La section suivante présente quelques solutions apportées pour améliorer la qualité du suivi.

7.2.2.3 Amélioration de la qualité du suivi

Le suivi de points, et par conséquent les applications de plus haut niveau qui s'y rattachent, peuvent être améliorés en rejetant tout simplement les points supposés mal suivis. Dans [Shi et Tomasi, 1994], le rejet s'effectue en fonction des résidus de convergence obtenus après minimisation de (7.5). Les points obtenant de forts résidus (comparativement à la majorité des points) sont rejetés. Les résidus attestent certes de la précision de la modélisation du mouvement mais également de la présence de bruit ou des variations d'illumination. Dans [Tommasini *et al.*, 1998b], l'auteur fait l'hypothèse selon laquelle la distribution des résidus de convergence d'un point correctement suivi satisfait une loi normale tout au long de la séquence d'images. Les points ne respectant pas cette contrainte sont rejetés.

Toutefois, ces approches se basent sur une vérification *a posteriori* du bon suivi des points, mais n'améliorent pas la tâche de suivi en elle-même, de sorte que des séquences d'images fortement bruitées peuvent amener à perdre un trop grand nombre de points et compromettre les applications de plus haut niveau. Une meilleure robustesse peut être obtenue en utilisant en ligne un estimateur robuste pour obtenir une mesure de confiance des points suivis par rapport au modèle de déplacement considéré [Odobez et Bouthemy, 1995]. Cette technique apporte également une meilleure robustesse vis-à-vis des occultations partielles de la fenêtre considérée. Citons également les travaux de [Arnaud *et al.*, 2004], où des trajectoires complexes sont appréhendées grâce à l'utilisation du filtrage particulaire. Dans cet article, l'association d'une

méthode de corrélation avec une approche différentielle permet d'assurer une bonne robustesse vis-à-vis des occultations totales.

Enfin, une bonne robustesse du suivi peut être obtenue en écartant *a priori* les mauvais candidats au suivi [Espiau *et al.*, 2002, Kermad *et al.*, 2002] à l'aide d'un critère de détection de ces points.

Ainsi, il existe deux types de méthodes de suivi de points : les approches corrélatives et les approches différentielles. Ces dernières permettent de calculer les déformations géométriques de la zone d'intérêt \mathcal{W} centrée autour du point à suivre. En outre, du fait qu'elles ne requièrent pas une recherche exhaustive des correspondances entre deux images de la séquence, elles s'avèrent très pertinentes dans le cadre de tâches de vision par ordinateur. Néanmoins, l'approche de suivi différentiel classique, décrite dans la section 7.2.2.1 repose sur le modèle de conservation de la luminance (section 6.2.1 page 131), or nous avons montré dans cette même section le caractère erroné de cette hypothèse dans le cas où les objets sont spéculaires ou encore lorsque l'éclairage n'est pas contrôlé. Nous consacrons la section qui suit aux approches différentielles de suivi de points proposant une certaine robustesse vis-à-vis de ces différents changements.

7.3 Robustesse du suivi de points différentiel vis-à-vis des variations d'illumination

Peu d'auteurs se sont intéressés à la robustesse des approches de suivi de points différentiel. Néanmoins, il existe une méthode permettant de compenser une modélisation affine des variations d'illumination décrite dans la section 6.2.4 et une seconde technique se basant sur une normalisation photométrique (comme dans [Tommasini *et al.*, 1999]). Nous montrons dans cette section que les deux approches se basent sur les mêmes hypothèses quant aux variations d'illumination.

7.3.1 Méthode de suivi compensant les changements affines d'illumination

Dans [Jin *et al.*, 2001] les auteurs proposent une méthode de suivi basée sur la compensation d'un modèle affine de changement d'illumination (voir la section (6.2.4) page 132) en calculant les paramètres de mouvement, d'intensité η et de contraste λ dans un même processus de minimisation :

$$\epsilon_2(\mathbf{A}, \lambda, \eta) = \sum_{m \in \mathcal{W}} (\lambda f(m) - g(\delta(m, \mathbf{A})) + \eta)^2 \quad (7.11)$$

Posons $\mathbf{B} = [\mathbf{A}^T, \lambda, \eta]^T$. En supposant une petite variation \mathbf{b} de \mathbf{B} autour d'un estimé $\hat{\mathbf{B}}$, on aboutit à un système linéaire en \mathbf{b} :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V}_s \mathbf{V}_s^T \right) \mathbf{b} = \sum_{m \in \mathcal{W}} \left(\hat{\lambda} f(m) - g(\delta(m, \hat{\mathbf{A}})) + \hat{\eta} \right) \mathbf{V}_s \quad (7.12)$$

où le vecteur \mathbf{V}_s vaut :

$$\mathbf{V}_s = [g_x, g_y, xg_x, xg_y, yg_x, yg_y, -f(m), -1]^T \quad (7.13)$$

Remarquons dès à présent que \mathbf{V}_s comprend un terme dépendant des luminances $f(m)$ sur la fenêtre d'intérêt \mathcal{W} . Cela implique premièrement que le conditionnement de la matrice à inverser $\sum_{m \in \mathcal{W}} \mathbf{V}_s \mathbf{V}_s^T$ dépend du contenu de \mathcal{W} . Ensuite, la dynamique de f étant généralement plus importante que celle de la dérivée, des problèmes de conditionnement peuvent survenir. Nous reviendrons de manière plus détaillée sur ce problème dans le cadre des expérimentations de la section 7.6.

La section suivante décrit une méthode de suivi de points basée sur une normalisation photométrique.

7.3.2 Suivi par normalisation photométrique

L'approche de suivi avec normalisation photométrique (voir les travaux de [Tommasini *et al.*, 1999] notamment) se base sur la minimisation du critère suivant :

$$\epsilon_3(\mathbf{A}) = \sum_{m \in \mathcal{W}} (\lambda f(m) - g(\delta(m, \mathbf{A})) + \eta)^2 \quad (7.14)$$

Il s'agit là encore d'un modèle affine d'illumination mais, contrairement à l'approche précédente où les paramètres λ et η étaient obtenus par minimisation, cette technique les mesure à chaque itération. Comme nous le verrons ci-après, ils sont exprimés par $\lambda = \frac{\sigma_g}{\sigma_f}$ et $\eta = -\frac{\sigma_g \mu_f}{\sigma_f} + \mu_g$, où μ_f et μ_g sont les moyennes de f et g calculées sur \mathcal{W} tandis que σ_f et σ_g correspondent aux écarts-type. Par conséquent, cette approche se fonde sur les mêmes hypothèses que la méthode décrite dans la section précédente, à savoir que λ et η sont supposés constants en tout point de \mathcal{W} . Or, d'après l'analyse de la section 6.2.4 (page 132), le modèle affine suppose que les paramètres du modèle de réflexion spéculaire sont constants sur \mathcal{W} (K_d , K_a , K_d , h_f et θ_i sont constants en tout point de \mathcal{W}). On montre que :

$$\begin{cases} \mu_f &= K_d \cos(\theta_i) \mu_a + K_a + h_f \\ \sigma_f &= K_d \cos(\theta_i) \sigma_a \end{cases} \quad (7.15)$$

où μ_a est la valeur moyenne de a sur \mathcal{W} et σ_a est son écart-type. Par ailleurs, les valeurs $f(m) - \mu_f$, $a(m) - \mu_a$ et $g(m') - \mu_g$ permettent d'éliminer le terme spéculaire h_f . Ensuite, les rapports qui suivent sont invariants à tout type de changement d'illumination :

$$\frac{f(m) - \mu_f}{\sigma_f} = \frac{a(m) - \mu_a}{\sigma_a} = \frac{g(m') - \mu_g}{\sigma_g}, \quad \forall m \in \mathcal{W} \quad (7.16)$$

puisque'ils ne dépendent plus que de la reflectance diffuse du matériau. On obtient finalement la relation suivante entre f et g :

$$g(m') = \frac{\sigma_g}{\sigma_f} f(m) - \frac{\sigma_g \mu_f}{\sigma_f} + \mu_g \quad (7.17)$$

En injectant (7.6) dans (7.14), on aboutit à un système linéaire en \mathbf{a} :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V}_n \mathbf{V}_n^T \right) \mathbf{a} = \sum_{m \in \mathcal{W}} \left(\lambda f(m) - g(\delta(m, \hat{\mathbf{A}})) + \eta \right) \mathbf{V}_n \quad (7.18)$$

où le vecteur \mathbf{V}_n vaut :

$$\mathbf{V}_n = [g_{x_n}, g_{y_n}, xg_{x_n}, xg_{y_n}, yg_{x_n}, yg_{y_n}]^T \quad (7.19)$$

avec $g_{x_n} = g_x/\sigma_g$ et $g_{y_n} = g_y/\sigma_g$. Néanmoins, la section 6.2.4 a mis en exergue que les hypothèses sur lesquelles se base cette technique ne sont pas réalistes dans certains cas de figure, en particulier lorsque \mathcal{W} est la projection d'une surface non plane de la scène. De plus, le calcul de la normalisation devient instable lorsque $\sigma_a \approx 0$.

Notons que des variantes de (7.14) pourraient être exploitées. Notamment, la différence entre les intensités de deux points m_0 et m_1 d'une même fenêtre d'étude \mathcal{W} ne dépend plus de la composante spéculaire de la réflexion :

$$f(m_0) - f(m_1) = [a(m_0) - a(m_1)] K_d \cos \theta_i \quad (7.20)$$

En considérant un troisième point $m_2 \in \mathcal{W}$, on montre facilement que le rapport de deux différences ne dépend plus que de la reflectance diffuse :

$$\frac{f(m_0) - f(m_1)}{f(m_0) - f(m_2)} = \frac{a(m_0) - a(m_1)}{a(m_0) - a(m_2)} \quad (7.21)$$

Ainsi tout rapport de différences d'intensité dans une fenêtre \mathcal{W} est un attribut invariant, qui ne dépend plus que des caractéristiques intrinsèques du matériau, la reflectance diffuse, et non plus des changements de direction d'observation et d'éclairage. Toutefois, dans le cas où $f(m_0) \simeq f(m_2)$, cette valeur est très bruitée, ce qui justifie pleinement l'utilisation de (7.14).

Ainsi, les approches de suivi différentielles compensant un modèle affine d'illumination ou proposant une normalisation photométrique se basent toutes les deux sur la constance des variations d'illumination sur \mathcal{W} . Les approches que nous proposons dans la suite ne font plus cette hypothèse, mais supposent que les changements d'illumination peuvent être approximés par des fonctions continues dans la zone d'intérêt considérée.

7.4 Méthode de suivi robuste aux variations de spécularité et aux changements d'éclairage

D'après l'étude effectuée dans le chapitre 6, la modélisation des variations de luminance entre f et g exprimée par $g(m) = f(m) + \phi(m)$ est la seule possibilité de prendre en compte les spécularités (voir la section 6.1.1 page 128) et correspond à l'un des deux modèles de variation d'éclairage possibles (voir la section 6.1.2). Dans la section 6.3.1, nous proposons une approximation locale de ϕ . Ainsi, nous calculons simultanément les paramètres de mouvement

\mathbf{A} et les trois paramètres de la modélisation photométrique $\mathbf{C} = (\alpha, \beta, \gamma)^T$ par minimisation de la fonction de coût suivante [Gouiffès *et al.*, 2005] :

$$\epsilon_4(\mathbf{A}, \mathbf{C}) = \sum_{m \in \mathcal{W}} (f(m) - g(\delta(m, \mathbf{A})) - \mathbf{U}^T \mathbf{C})^2 \quad (7.22)$$

où $\mathbf{U} = (x - x_p, y - y_p, 1)^T$. Posons $\mathbf{D} = [\mathbf{A}^T, \mathbf{C}^T]^T$. En supposant une petite variation \mathbf{d} de \mathbf{D} autour d'un estimé $\hat{\mathbf{D}}$, (7.22) peut être linéarisé en \mathbf{d} si bien que l'algorithme de suivi consiste à minimiser le critère suivant :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V}_p \mathbf{V}_p^T \right) \mathbf{d} = \sum_{m \in \mathcal{W}} (f(m) - g(\delta(m, \hat{\mathbf{A}})) - \mathbf{U}^T \hat{\mathbf{C}}) \mathbf{V}_p \quad (7.23)$$

où le vecteur \mathbf{V}_p est exprimé par :

$$\mathbf{V}_p = [g_x, g_y, xg_x, xg_y, yg_x, yg_y, x - x_p, y - y_p, 1]^T \quad (7.24)$$

Contrairement aux approches abordées dans la section 7.3, les variations d'illumination ne sont pas considérées spatialement constantes. Néanmoins, lorsque des variations d'éclairage se produisent, le modèle d'illumination suppose que $a(p)$, relatif à la réflectance diffuse, peut être approximé par une série de Taylor au 1^{er} ordre (voir l'expression de ϕ donnée par (6.18)). Toutefois, cette limitation n'est pas forcément avérée dans le cas du suivi de points, où de petites fenêtres d'intérêt sont considérées. En revanche, ce modèle peut être limitatif pour de plus larges fenêtres, c'est pourquoi nous proposons ci-dessous une approche de suivi plus adaptée à ce cas de figure.

7.5 Extension au suivi de motifs plus larges

De part la bonne fiabilité du modèle de mouvement considéré dans les approches de suivi différentielles, ces méthodes peuvent être étendues au cas de plus larges zones de l'image que dans le cas du suivi de points. C'est le cas des travaux [Hager et Belhumeur, 1998] et [Jurie et Dhome, 2002]. Notons que les variations d'illumination sont prises en compte dans [Hager et Belhumeur, 1998], où les auteurs utilisent une base d'images acquises hors ligne suivant différentes configurations d'illumination, permettant de décrire l'ensemble des différents changements d'apparence possibles. Cette méthode est certes efficace mais la nécessité d'effectuer un apprentissage peut être considérée limitative.

Dans la section 6.3.2, nous avons défini un modèle photométrique complet, compensant les variations spatiales des changements de spécularité et d'éclairage. Nous utilisons ce modèle dans le cadre du suivi de zones d'intérêt. Le paramètre de mouvement \mathbf{A} et les paramètres de réflexion $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^T$ et $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)^T$ sont obtenus par minimisation du critère suivant :

$$\epsilon_5(\mathbf{A}, \boldsymbol{\lambda}, \boldsymbol{\eta}) = \sum_{m \in \mathcal{W}} (\mathbf{U}^T \boldsymbol{\eta} f(m) - g(\delta(m, \mathbf{A})) - \mathbf{U}^T \boldsymbol{\lambda})^2 \quad (7.25)$$

où $\mathbf{U} = (x - x_p, y - y_p, 1)^T$. On pose $\mathbf{E} = [\mathbf{A}^T, \boldsymbol{\lambda}^T, \boldsymbol{\eta}^T]^T$. Ici aussi, en considérant une petite variation e de \mathbf{E} autour d'un estimé $\hat{\mathbf{E}}$, le système (7.25) peut être linéarisé, comme dans 7.2.2.1, conduisant à la minimisation du critère suivant :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V}_m \mathbf{V}_m^T \right) e = \sum_{m \in \mathcal{W}} \left(\mathbf{U}^T \hat{\boldsymbol{\lambda}} f(m) - g(\delta(m, \hat{\mathbf{A}})) - \mathbf{U}^T \hat{\boldsymbol{\eta}} \right) \mathbf{V}_m \quad (7.26)$$

Le vecteur \mathbf{V}_m est exprimé par :

$$\mathbf{V}_m = [g_x, g_y, xg_x, xg_y, yg_x, yg_y, (x - x_p)f(m), (y - y_p)f(m), f(m), x - x_p, y - y_p, 1]^T \quad (7.27)$$

Intéressons-nous à la matrice $\sum_{m \in \mathcal{W}} \mathbf{V}_m \mathbf{V}_m^T$. Elle fait intervenir des termes dépendant de l'image et comprend des valeurs très différentes, telles que des intensités et des différences d'intensités. De ce fait, comme l'approche décrite dans (7.3.1), elle peut être mal conditionnée. De plus, le nombre de paramètres à calculer est plus important puisque douze paramètres sont nécessaires dans le cas d'un modèle affine de mouvement. De ce fait, l'utilisation de trop petites fenêtres \mathcal{W} peut nuire à la robustesse du suivi. C'est pourquoi l'utilisation de ce modèle n'est pas justifiée pour assurer le suivi de points d'intérêt, en comparaison avec la modélisation $g(m) = f(m) + \phi(m)$. Nous la destinons plutôt à une utilisation sur des fenêtres plus grandes, exploitées par exemple en calcul de mouvement ou en suivi d'objets.

La section qui suit permet de vérifier, par le biais de simulations, l'apport des approches de suivi proposées, décrites dans les sections 7.4 et 7.5.

7.6 Comparaison des méthodes de suivi vis-à-vis des variations d'illumination

Cette section présente des résultats de simulation et des résultats expérimentaux. Les différents algorithmes de suivi présentés dans les sections 7.2.2.1, 7.3.1, 7.3.2, 7.4 et 7.5 sont comparés sur des séquences d'images montrant des déformations géométriques et des variations photométriques. Les points d'intérêt sont sélectionnés par le détecteur de Harris [Harris et Stephens, 1988], décrit dans la section 7.1.2 et les points sélectionnés sont les mêmes pour chacune des approches. Les méthodes de suivi sont basées sur le calcul d'un modèle de mouvement affine entre la première image et l'image courante (voir la section 7.2.2.2). Les algorithmes rejettent les mauvais points en se basant sur l'analyse de la convergence des résidus $\epsilon_i, i = 1 \dots 5$. Ainsi, un point p est rejeté du processus de suivi dès que les résidus deviennent plus élevés que $\mathcal{S}_{conv} = \mathcal{N}^2 E_{moy}^2$, où E_{moy} est la variation d'intensité tolérée sur \mathcal{W} entre f et g et \mathcal{N} est la largeur de la fenêtre \mathcal{W} . Dans le cadre de cette étude comparative, nous avons fixé E_{moy} à 15. Usuellement, le choix de la taille \mathcal{N} est basé sur un compromis entre robustesse par rapport au bruit, temps de calcul et concordance vis-à-vis des hypothèses sur lesquelles les techniques de suivi sont basées, telle que la planarité de la surface. Elle dépend également de l'application visée. À titre de comparaison, nous considérons ici plusieurs tailles différentes. Nous étendons même l'étude au cas de larges motifs (de taille $\mathcal{N} = 35$ et $\mathcal{N} = 151$), de manière à évaluer la validité des différentes modélisations photométriques dans ces cas de figure.

Afin de comparer l'efficacité des méthodes de suivi, plusieurs critères sont calculés :

- la *robustesse* du suivi, c'est-à-dire le nombre de points correctement suivis durant toute la séquence ;
- la *précision* du suivi : nous calculons la moyenne des résidus pour l'ensemble des points qui ont été suivis durant toute la séquence par la méthode considérée. Ce critère nous informe quant à la validité du modèle photométrique utilisé. Plus les résidus sont faibles, au mieux les variations d'illumination sont compensées.

En outre, les résultats de simulation permettent de comparer l'erreur d'estimation du positionnement des points des approches de suivi puisque l'on dispose de la position réelle des points à suivre. Nous montrons également l'évolution des paramètres photométriques calculés avec les modèles proposés. Dans toutes ces expérimentations, nous considérons les notations et les codes couleur (quant aux courbes tracées) suivants :

	C : méthode classique de suivi de points (section 7.2.2.1)
	S : méthode calculant les paramètres affines d'illumination (section 7.3.1)
	N : approche avec normalisation photométrique (section 7.3.2)
	P_3 : approche de suivi de points proposée (section 7.4)
	P_6 : l'approche dédiée aux motifs plus larges (section 7.5)

7.6.1 Résultats de simulation

L'évaluation des performances des méthodes n'est pas une tâche aisée. Elle passe par l'observation conjointe de plusieurs critères : le nombre de points perdus, les résidus de convergence moyens et l'erreur de position (la distance euclidienne) entre la position estimée des points suivis et leur position réelle mesurée à l'aide du logiciel de simulation. De manière à quantifier les performances des différentes approches, nous proposons le critère suivant :

$$\mathcal{C} = (\max_{Res} + \epsilon)(\max_{Err} + \epsilon)(NbPts + \epsilon) \quad (7.28)$$

\max_{Res} correspond à la valeur maximale des résidus moyens obtenus au cours de la séquence et \max_{Err} à l'erreur maximale. Enfin, $NbPts$ est le nombre de points perdus par la méthode. Le paramètre ϵ permet simplement que \mathcal{C} ne soit pas annulé dans le cas où l'un des paramètres \max_{Res} , \max_{Err} et $NbPts$ est nul. On prend ici $\epsilon = 0.1$. Nous considérons un objet plan puis un objet de forme cylindrique de rayon 12 cm, en analysant d'une part le cas où des spéularités et un déplacement de la source sont produits, d'autre part le cas où spéularités, variations d'intensité d'éclairage et de position de la source sont simultanément provoquées. La caméra est située à une distance $Z_P = 1$ m de l'objet et la source est situé à $S_z = 50$ cm. On considère un déplacement $dS_X = 30$ cm $dS_Y = 30$ cm.

7.6.1.1 Objet plan

La figure 7.2 représente trois images d'une séquence simulée représentant un objet plan soumis à des variations de position d'éclairage et de spéularités. 13 points sont sélectionnés dans la première image. Enfin, trois tailles de fenêtre sont considérées ($\mathcal{N} = 7, 35$ et 151).

Déplacement de la source et spécularités

◦ $\mathcal{N} = 7$ (suivi de points). Pour cette taille de fenêtre, les méthodes S et P_3 ne perdent aucun point tandis que N en perd 2 et C en perd 8. Enfin, l'approche P_6 n'est pas considérée puisqu'elle ne converge pas correctement. Les évolutions des résidus de convergence obtenus pour $\mathcal{N} = 7$ sont décrites par la figure 7.3a. Notons que l'approche classique C n'est pas prise en compte sur cette figure du fait que les résidus sont bien plus élevés que ceux des autres approches.

L'évolution des erreurs moyennes entre la position estimée des points et leur position réelle est montrée sur la figure 7.3b. On constate pour les approches C , N et S des chutes brutales des erreurs, qui sont dues à la perte des points mal suivis. En outre, lorsqu'une technique perd un point pour lequel les résidus sont élevés, la moyenne des résidus se voit généralement diminuée (voir la figure 7.3a).

Les résultats obtenus amènent à deux constats intéressants. Tout d'abord, P_3 s'avère plus précise que les autres approches. Les différences de résidus ne sont pas significatives, mais l'erreur de positionnement obtenue par la méthode S est très élevée. Cette approche a abouti au calcul de paramètres de mouvement erronés. Les résidus moyens obtenus par N sont plus faibles que ceux de P_3 , mais ce constat est contrebalancé par le fait que le nombre de points suivis est plus faible. Finalement, au sens du critère \mathcal{C} (tableau 7.1), P_3 obtient les meilleures performances (\mathcal{C} plus faibles).

◦ $\mathcal{N} \geq 35$ (suivi de motifs). L'approche classique perd l'ensemble des points, alors que les autres approches les suivent tous. D'après la figure 7.3c, La méthode P_6 obtient de bien plus faibles résidus que les autres techniques. Ceci en fait également l'approche la plus performante au sens du critère \mathcal{C} (voir le tableau 7.1). Nous n'avons pas affiché les erreurs de positionnement obtenues pour cette taille de fenêtre, puisqu'elles s'avèrent très faibles et fort comparables d'une approche à l'autre.

◦ $\mathcal{N} = 151$ (suivi de motifs larges). Une seule fenêtre centrée sur l'objet est prise en compte. Elle est suivie correctement par chacune des approches (N , S , P_3 , P_6), mais pas par l'approche classique. D'après la figure 7.3d, P_6 aboutit à des résidus plus faibles que les autres techniques. Finalement, d'après le tableau 7.1, elle obtient une valeur de critère \mathcal{C} plus faible. Elle correspond à une modélisation plus appropriée des changements de luminance.

Intéressons-nous à présent au cas de variations d'éclairage.

Spécularités, variations d'intensité et de position d'éclairage

Considérons à présent que des variations d'intensité d'éclairage sont provoquées en plus des variations de spécularité et de position de la source. L'intensité des éclairages direct et ambiant (paramètres K_d , K_s , K_a du modèle de Phong, décrit par l'expression (1.10)), varie aléatoirement de 10% autour de la valeur initiale.

◦ $\mathcal{N} = 7$ (suivi de points). Au cours de cette séquence, la méthode P_3 est la seule des approches à ne perdre aucun point parmi les 13 sélectionnés, alors que S en perd 1 et N en

perd 3. Enfin l'approche classique C perd 12 points. Là encore, l'approche P_6 n'a pas convergé correctement. Les résidus de convergence, montrés sur la figure 7.4a confirment les bonnes performances de P_3 , pour laquelle les résidus atteignent des valeurs moins élevées que les autres techniques¹. La technique N obtient quant à elle les valeurs de résidus les plus élevées. En outre, les erreurs de positionnement représentées sur la figure 7.4b attestent que cette approche apporte une bonne précision de suivi par rapport aux autres techniques. Néanmoins, elle atteint une valeur élevée de 5 pixels d'erreur, qui s'explique concrètement par le mauvais suivi de l'un des points. Les variations d'intensité d'éclairage sont importantes, ce qui correspond aux conditions les moins favorables de validité de P_3 . Dans ce cas, d'après la relation (6.18) page 129, les paramètres photométriques calculés par P_3 doivent compenser une fonction dépendant de la réflectance. Notons ensuite que d'après la figure 7.4b, les erreurs engendrées par C et S sont très importantes. La chute de l'erreur de C à l'itération 63 correspond à la perte du point mal suivi. L'apparition de telles erreurs est due à la valeur choisie de E_{moy} . Fixée ici à 15 de manière à comparer les différentes approches, cette valeur pourrait être diminuée de sorte d'éliminer plus strictement les points mal suivis.

Néanmoins, les valeurs de critère \mathcal{C} (tableau 7.1) soulignent finalement les meilleures performances de P_3 par rapport aux autres approches.

- Suivi de zones de l'image ($\mathcal{N} = 35$ et $\mathcal{N} = 151$). D'après l'évolution des résidus de convergence des figures 7.4c et 7.4d, respectivement pour $\mathcal{N} = 35$ et $\mathcal{N} = 151$, l'approche P_6 s'avère plus précise dans les deux cas. Les valeurs de critères d'évaluation, données dans le tableau 7.1, confirment là encore ses meilleures performances.

Étudions à présent le cas d'un objet cylindrique.

7.6.1.2 Objet cylindrique

La figure 7.5 représente trois images d'une séquence simulée (de 100 images) représentant un objet de forme cylindrique soumis à des variations d'intensité, de position d'éclairage et de spécularités. Notons que des variations d'illumination très importantes sont provoquées entre les instants 30 et 70 (voir la figure 7.5b).

Déplacement de la source et spécularités

- $\mathcal{N} = 7$ (suivi de points). Pour cette taille de fenêtre, la méthode P_3 ne perd aucun point, la méthode S en perd 2 et enfin la technique N en perd 5, sachant que 7 points ont été sélectionnés initialement. En revanche, l'approche classique perd l'ensemble des points. Les résidus de convergence de la figure 7.5a attestent que les deux points suivis par la méthode normalisée sont très bien suivis, les résidus étant très faibles. L'approche P_3 obtient des résidus plus élevés, mais qui sont très corrects sur l'ensemble des 7 points suivis. Enfin, la figure 7.5b, relative à l'erreur de position, montre que les approches S et C ont mal convergé et ont abouti à de mauvais paramètres de mouvement.

¹Là encore, les résidus obtenus par l'approche classique ne sont pas présentés du fait de la valeur très élevée des résidus obtenus.

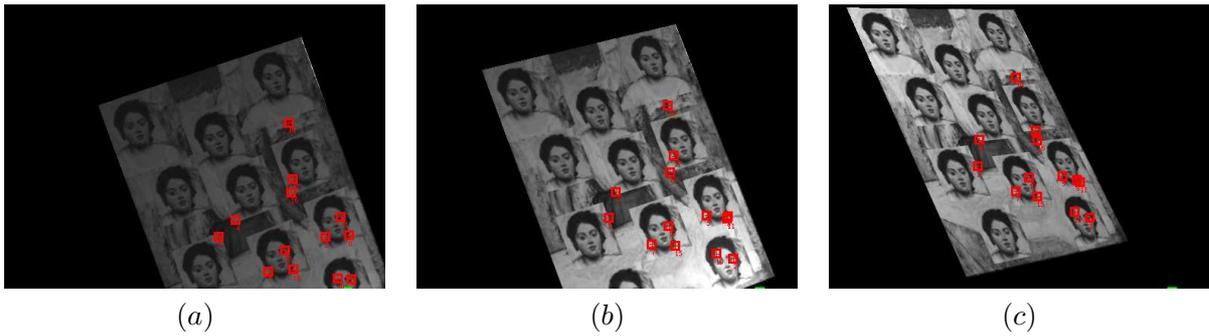


FIG. 7.2 – Images de la séquence simulée d'un objet plan soumis à des variations d'illumination. (a) Image 1. (b) Image 25. (c) Image 99.

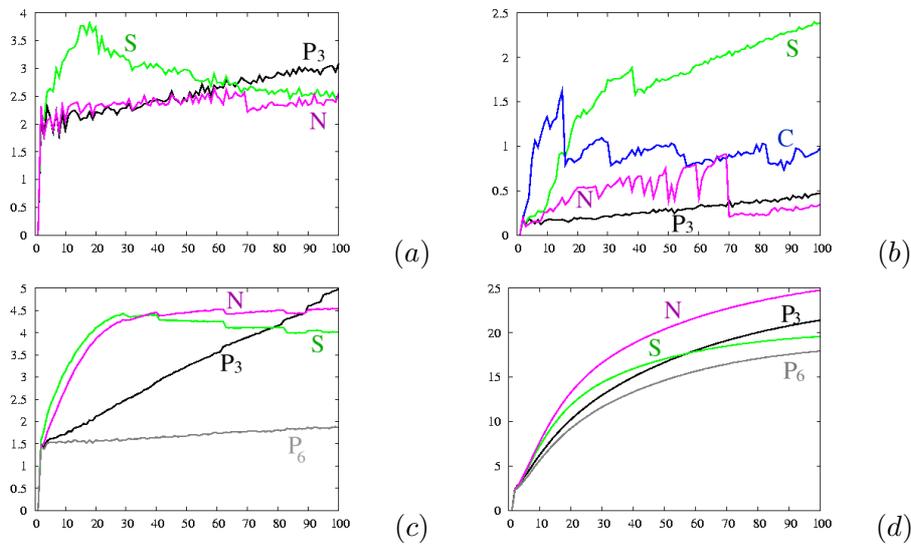


FIG. 7.3 – Objet plan, variations de spécularités. (a) Résidus de convergence obtenus pour $\mathcal{N} = 7$. (b) Moyenne des erreurs de position (en pixels) des points suivis pour $\mathcal{N} = 7$ (en pixels). (c) et (d) Résidus de convergence obtenus respectivement pour $\mathcal{N} = 35$ et $\mathcal{N} = 151$.

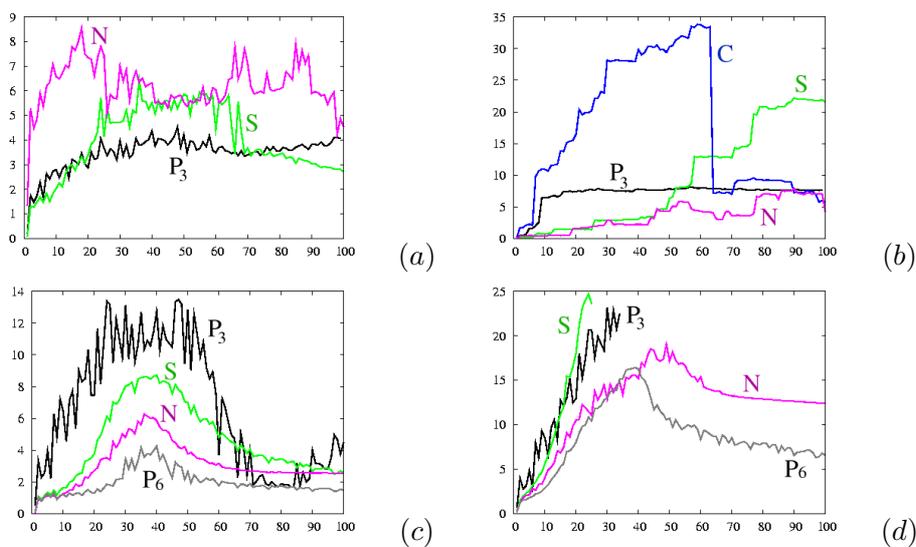


FIG. 7.4 – Objet plan, variations de spécularités et d'éclairage. (a) Résidus de convergence obtenus pour $\mathcal{N} = 7$. (b) Moyenne des erreurs de position (en pixels) des points suivis pour $\mathcal{N} = 7$. (c) et (d) Résidus de convergence obtenus pour $\mathcal{N} = 35$ et $\mathcal{N} = 151$.

◦ $\mathcal{N} = 35$ (suivi de motifs). P_3 et P_6 sont les seules approches permettant de suivre tous les points tout au long de la séquence. L'approche classique les perd tous, tandis que les méthodes S et N en perdent 5. D'après les résidus de convergence obtenus (voir la figure 7.6c), P_6 permet finalement de compenser de manière plus significative les variations d'illumination, comparativement aux autres techniques, dont les résidus sont plus élevés. Les critères \mathcal{C} (tableau 7.1) viennent confirmer ce constat, puisque P_6 obtient des valeurs plus faibles.

◦ $\mathcal{N} = 121$ (suivi de motifs larges). Seules les approches P_3 et P_6 permettent de suivre la large zone de l'image sélectionnée. Les autres techniques perdent le motif. D'après les résidus obtenus (figure 7.6d), P_6 permet une meilleure compensation des variations d'illumination tout au long de la séquence, et en particulier entre les itérations 30 et 70, où se produisent les variations d'illumination les plus notables. Néanmoins, au sens du critère \mathcal{C} , les deux approches sont très comparables.

Spécularités, variations d'intensité et de position d'éclairage

Les variations d'éclairage provoquées sont les mêmes que celles qui ont été décrites dans le cas de la simulation sur l'objet plan (voir la section 7.6.1.1).

◦ $\mathcal{N} = 7$ (suivi de points). Pour cette taille de fenêtre, la méthode P_3 ne perd aucun point alors que les méthodes S et N en perdent 4. L'approche classique perd quant à elle l'ensemble des points. D'après l'évolution des résidus de convergence (voir la figure 7.7a), la moyenne des résidus obtenue avec la méthode P_3 sur les 7 points suivis atteint des valeurs plus faibles que celle obtenue par N et S sur les 3 points qu'elles suivent. Notons également que les résidus sont plus bruités que pour la séquence précédente, du fait des variations aléatoires de l'intensité d'éclairage. Comme pour la simulation précédente, les erreurs mesurées (voir la figure 7.7b) pour l'approche S sont très importantes. Elles ont très mal convergé pour certains points. Les approches P_3 et N obtiennent quant à elles une erreur de suivi très faible tout au long de la séquence.

◦ $\mathcal{N} = 35$ (suivi de motifs larges). Les résultats de suivi sont approximativement similaires à ceux obtenus pour une même taille de fenêtre à la simulation précédente. Notamment les nombres de points suivis restent les mêmes. Par contre, les résidus sont plus élevés que précédemment pour l'ensemble des approches. Là encore l'approche P_6 obtient les résidus les plus faibles. Signalons également que S est plus précise dans le cas de variations d'intensité d'éclairage (figure 7.7c) que dans le cas de changements de spécularité (figure 7.6c). Enfin, les approches N et P_3 obtiennent des résidus très comparables.

◦ $\mathcal{N} = 121$ (suivi de motifs plus étendus). Les conclusions qui peuvent être tirées ici sont les mêmes que dans le cas de figure précédent, où seules des spécularités étaient traitées. L'approche P_6 est là encore la méthode la plus performante.

7.6.1.3 Bilan des résultats de simulation

Quels que soient le type de variation d'illumination et la forme de l'objet pris en compte, l'approche P_3 proposée pour le suivi de points ($\mathcal{N} = 7$) s'avère plus performante que les autres

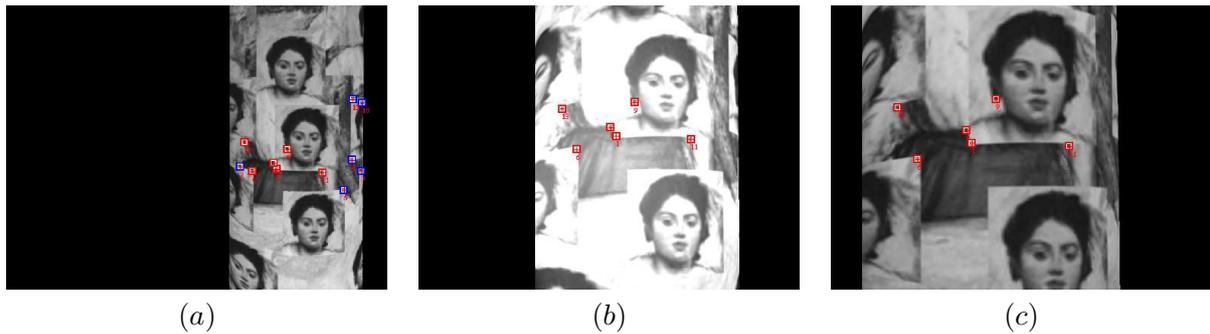


FIG. 7.5 – Images de la séquence simulée d'un objet de forme quadratique soumis à des variations d'illumination. (a) Image 1. (b) Image 25. (c) Image 75. (d) Image 99.

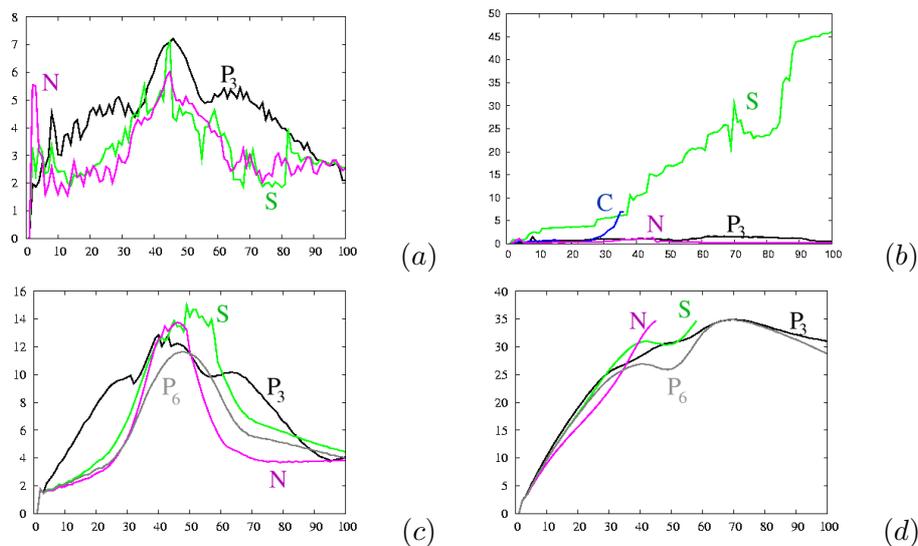


FIG. 7.6 – Objet de forme cylindrique, variations de spécularités. (a) Résidus de convergence obtenus pour $\mathcal{N} = 7$. (b) Moyenne des erreurs de position des points suivis pour $\mathcal{N} = 7$ (en pixels). (c) et (d) Résidus pour 35 et 151.

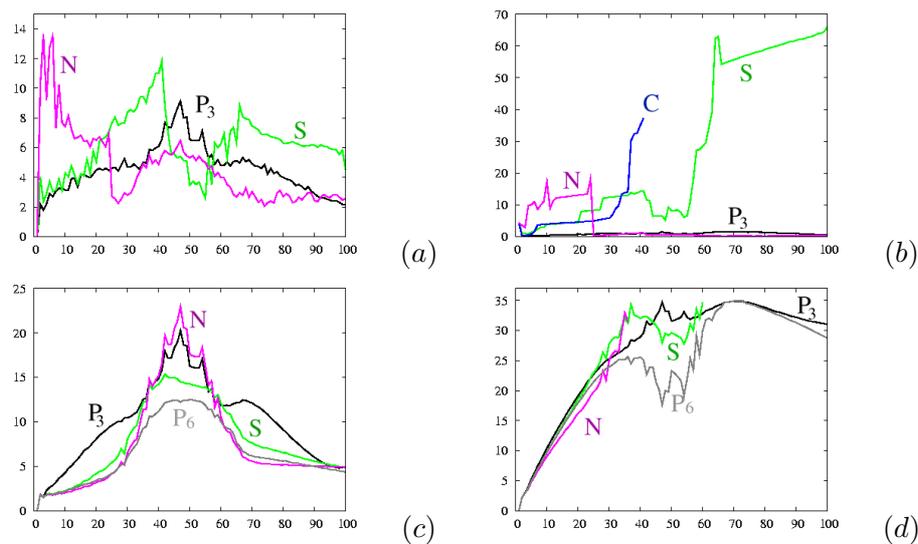


FIG. 7.7 – Objet de forme cylindrique, variations de spécularités et d'éclairage. (a) Résidus de convergence obtenus pour $\mathcal{N} = 7$. (b) Moyenne des erreurs de position des points suivis pour $\mathcal{N} = 7$ (en pixels). (c) et (d) Résidus de convergence obtenus pour 35 et 151.

TAB. 7.1 – Valeurs du critère C obtenues par les approches C , N , S , P_3 et P_6 dans les différentes simulations effectuées, selon que l'objet soit plan ou cylindrique, et que l'on considère des variations de spécularités ou des changements d'éclairage. Une valeur faible de C correspond à une bonne performance en terme de robustesse, de précision du suivi et de précision de la modélisation photométrique. Les valeurs en caractères gras correspondent aux valeurs les plus faibles pour une simulation donnée.

Type d'objet	Variation	\mathcal{N}	Méthode				
			C	N	S	P_3	P_6
Plan	Spécularités	7	374.00	5.60	0.84	0.18	-
		35	-	0.11	0.11	0.09	0.04
		151	-	20.55	16.29	17.57	14.78
	Intensité	7	3469	83.43	154.65	3.48	-
		35	75.88	0.16	0.21	37.79	0.07
		151	-	18.72	-	-	16.30
Cylindre	Spécularités	7	-	43	668.50	1.23	-
		35	-	70.03	10.67	38.97	0.92
		151	-	-	-	12.11	12.10
	Intensité	7	-	969	3303	1.54	-
		35	-	2.20	1.63	2.38	1.33
		151	-	-	-	12.28	12.13

approches en terme de robustesse (nombre de points suivis), tout en assurant une très bonne précision du suivi (des résidus de convergence faibles). Les erreurs de position sont également très satisfaisantes en comparaison par rapport aux techniques existantes. Néanmoins certains points mal suivis pourraient être plus strictement rejetés par une valeur de E_{moy} plus faible. Les valeurs du critère C obtenues (tableau 7.1), toujours plus faibles pour cette approche dans le cas $\mathcal{N} = 7$, viennent confirmer ce constat.

L'approche P_6 proposée pour le suivi de larges motifs s'avère toujours la plus performante (pour $\mathcal{N} \geq 35$) que les autres techniques, en terme de robustesse et en terme de précision.

D'après le tableau 7.1, pour $\mathcal{N} = 7$, les méthodes N et S qui prennent en compte un modèle d'illumination affine, obtiennent des valeurs de critères plus significativement importantes par rapport à P_3 dans le cas où l'objet est non plan (le cylindre) que dans le cas d'un objet plan. Cela tend à montrer que ces approches sont mieux adaptées dès lors que la surface observée est plane.

Dans la section suivante, des expérimentations sur séquences réelles permettent de valider notre approche.

7.6.2 Résultats d'expérimentation

Afin d'analyser correctement le comportement des différentes approches, chacune d'entre elles est testée tout d'abord sur des scènes où apparaissent des spécularités. Ensuite, les variations d'éclairage sont également prises en compte.

7.6.2.1 Occurrence de spéularités

Les deux premières séquences appelées *Livre* (200 images) et *Cylindre* (150 images) (respectivement figure 7.8a et figure 7.9a) représentent des objets spéculaires, respectivement de forme plane et cylindrique. Dans chaque cas, la scène et l'éclairage sont fixes mais la caméra est déplacée, ce qui entraîne des variations de spéularité à la surface des objets. Remarquons sur les figures 7.8a que des spéularités disparaissent au cours du mouvement, tandis que d'autres apparaissent à d'autres endroits sur la surface de l'objet. Dans le cas du cylindre (figure 7.9a), deux sources d'éclairage sont considérées, faisant apparaître deux zones fortement saturées. Au cours du mouvement de la caméra, les variations de spéularité sont importantes au voisinage de ces deux zones. Dans chaque cas, la séquence est jouée de la première à la dernière image, puis de la dernière à la première, de manière à vérifier la symétrie des paramètres photométriques calculés. Nous analysons successivement la robustesse, la précision des modélisations, puis nous analysons l'évolution des paramètres photométriques calculés. Les temps de calcul seront également pris en compte.

Robustesse. Un nombre total de 97 points est sélectionné initialement dans la séquence *Livre* et 137 dans la séquence *Cylindre*. Les tableaux 7.2a et 7.2b regroupent respectivement le pourcentage de points correctement suivis (par rapport aux points qui ne sont pas occultés ou qui ne sortent pas de l'image) en fonction de \mathcal{N} , pour chacune des techniques de suivi considérées. En ce qui concerne la séquence montrant un objet plan (séquence *Livre*), P_3 suit un plus grand nombre de points que les autres approches pour des tailles de fenêtre $\mathcal{N} \leq 15$. Dans le cas où un objet cylindrique est considéré (séquence *Cylindre*) elle demeure la plus robuste jusqu'à $\mathcal{N} = 25$. Par conséquent, l'apport de cette technique est plus significatif par rapport aux approches existantes dès lors que la surface mise en jeu est non plane. C'est en effet dans ce cas que les changements d'illumination sont susceptibles d'être différents en chaque point d'une fenêtre d'intérêt. Ce résultat valide l'étude théorique des modélisations, effectuée dans la section 6.4 page 136. Par contre, pour les tailles de fenêtre plus larges, la méthode P_6 devient plus performante dans le sens où elle permet de suivre un plus grand nombre de points. Par contre, cette technique ne converge pas pour de petites tailles de fenêtres, le nombre de paramètres à approximer étant trop important par rapport aux luminances disponibles, mais également par rapport au bruit. Notons ensuite que dans les deux séquences, S n'apporte pas une meilleure robustesse que C pour $\mathcal{N} \leq 13$. Par contre, elle devient plus performante que C pour des tailles de fenêtres plus grandes. Par conséquent, cette approche ne paraît pas adaptée aux séquences montrant des variations de spéularité dès lors que la fenêtre d'intérêt est trop restreinte. Enfin, N s'avère toujours plus robuste que l'approche classique C dans le cas de l'objet plan (séquence *Livre*). Par contre, lorsqu'un objet de forme non plane est considéré (séquence *Cylindre*), elle est moins performante que C pour $\mathcal{N} < 25$. Cette approche apparaît donc plus adaptée aux images présentant des surfaces planes.

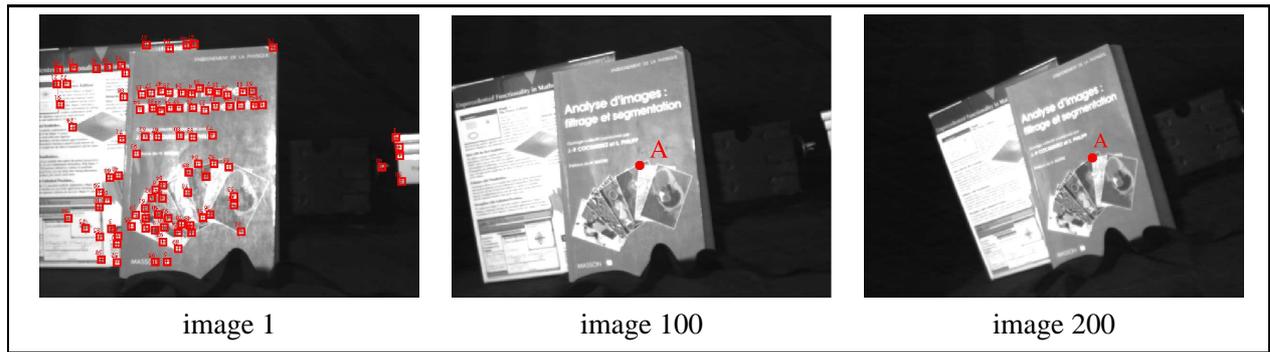
Précision des modèles. Les figures 7.8b et 7.9b comparent les résidus moyens obtenus par les différentes approches respectivement au cours des séquences *Livre* et *Cylindre*, pour de petites fenêtres d'intérêt ($\mathcal{N} = 9$). Dans le cas d'un objet plan (séquence *Livre*), N obtient de plus faibles résidus moyens que P_3 (figure 7.8b). Néanmoins, la moyenne des résidus est calculée sur 68 points dans le cas de P_3 et 33 dans le cas de N . Afin d'affiner la comparaison, la figure 7.8c compare les résidus moyens obtenus sur les mêmes points, ceux qui ont été suivis à la

fois par N et P_3 . Elle montre que P_3 obtient des résidus plus faibles, ce qui prouve la pertinence du modèle photométrique considéré dans le cas d'apparitions de spécularités. Dans le cas d'un objet cylindrique (séquence *Cylindre*), P_3 obtient de meilleurs résidus que les autres approches. Les résidus de la méthode N s'avèrent très élevés au début de la séquence (avant l'itération 50) et diminuent ensuite lorsque les points mal suivis ont été perdus. Ensuite, les figures 7.8d et 7.9c montrent les résidus de convergence moyens obtenus avec une taille de fenêtre $\mathcal{N}=35$. P_6 s'avère à chaque fois plus précise que toutes les autres approches. La moyenne de ses résidus (calculée sur 71 points suivis pour la séquence *Livre* et 105 points pour la méthode *Cylindre*) est inférieure à la moyenne des résidus de P_3 (calculée respectivement sur 65 et sur 92 points). Cela signifie que non seulement P_6 suit un plus grand nombre de points mais également que la modélisation photométrique est plus adaptée.

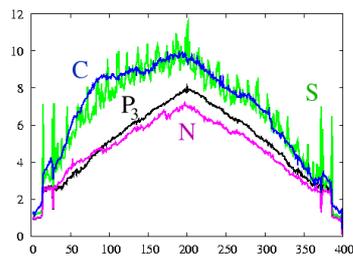
Paramètres photométriques. Pour analyser les variations d'illumination, nous avons sélectionné le point A visible sur la figure 7.8a. Il est situé dans une zone de fortes spécularités. La première rangée d'images de la figure 7.8e montre les luminances de la zone \mathcal{W} centrée autour du point A tandis que la seconde rangée correspond à cette même fenêtre après correction photométrique par P_6 , à différentes itérations, et pour $\mathcal{N} = 35$. Notons en particulier sur la première ligne que la dernière image est moins lumineuse que les précédentes, tandis que l'utilisation du modèle d'illumination a permis de compenser ces variations (la correction géométrique n'est pas affichée ici). Les variations spatiales du modèle d'illumination (ηU^T et λU^T avec $U^T = (x - x_p, y - y_p, 1)$) sont représentées sur la figure 7.8f, le niveau d'intensité étant proportionnel à la correction. Ils prouvent l'importance des variations d'illumination provoquées et montrent que ces changements ne sont pas constants sur \mathcal{W} . Pour comprendre l'évolution temporelle des modèles photométriques, analysons la figure 7.8f, qui trace les paramètres λ_i et η_i pour $i = 1..3$. La symétrie des courbes témoigne dans chaque cas que pour deux occurrences d'une même image (rappelons que la séquence est jouée de la première image à la dernière puis de la dernière à la première), les paramètres photométriques calculés sont les mêmes, ce qui tend à prouver la bonne convergence de l'approche. On notera toutefois que, malgré le fait que les variations de spécularité sont théoriquement modélisées par η , les paramètres λ compensent également une partie de ces variations. Dans le cas de la séquence *Cylindre*, l'évolution des paramètres associés au point A (visible sur l'image 7.9a) est montrée sur la figure 7.9d. Là encore, les courbes obtenues prouvent la bonne convergence de l'approche. Les paramètres λ compensent également une partie des variations spéculaires.

Temps de calcul. Considérons un point suivi par toutes les approches. Les temps de calcul des différentes méthodes sont regroupés dans le tableau 7.3, pour $\mathcal{N}=9, 15$ et 35. Remarquons que N et P_6 sont les approches les plus coûteuses en temps de calcul (soit du fait du calcul de la normalisation photométrique, soit à cause du grand nombre de paramètres à approximer). Ces fortes valeurs s'expliquent également par la mauvaise convergence de ces techniques dans le cas de petites fenêtres d'intérêt. L'algorithme étant itératif, elles requièrent un nombre d'itérations plus important pour aboutir à la convergence. Pour $\mathcal{N}=9$ et 15, les techniques C, S et P_3 obtiennent des temps de calcul très comparables.

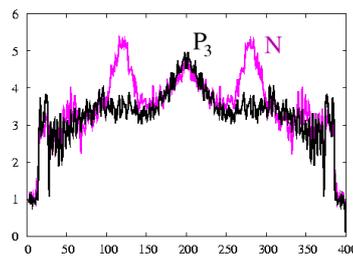
Jusqu'à présent, les tailles de motifs considérées dans ces expérimentations étaient relativement restreintes ($\mathcal{N} \leq 35$). Nous proposons une expérimentation supplémentaire permettant de comparer les approches dans le cadre du suivi d'un motif plus large, de taille $\mathcal{N} = 151$.



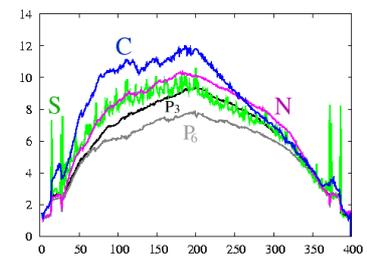
(a)



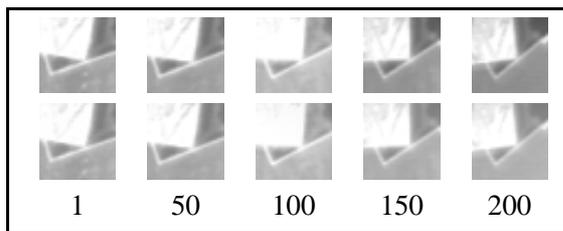
(b) $\mathcal{N} = 9$



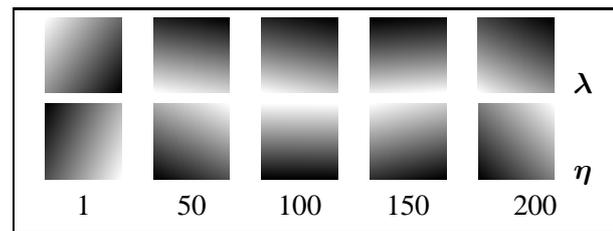
(c) $\mathcal{N} = 9$



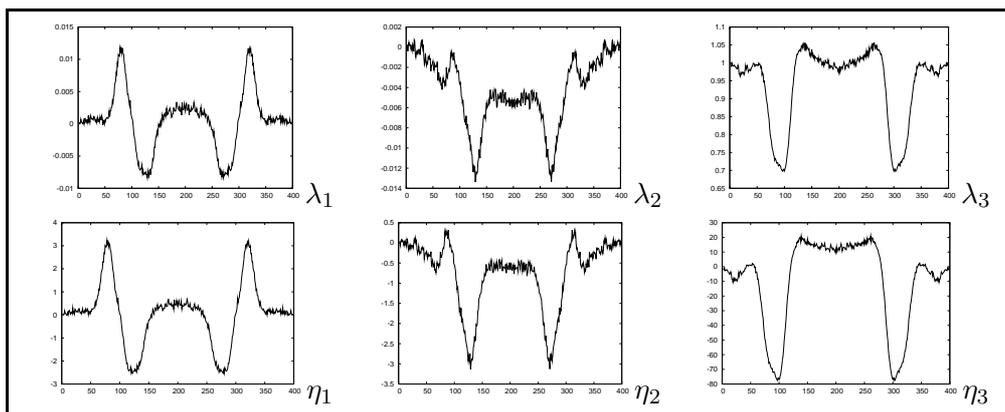
(d) $\mathcal{N} = 35$



(e)



(f)



(g)

FIG. 7.8 – Expérimentation Livre (97 points sont sélectionnés). (a) Images de la séquence aux itérations 1, 100 et 200. (b) Résidus de convergence moyens pour $\mathcal{N}=9$. (c) Comparaison des résidus de convergence obtenus sur les points suivis à la fois par N et P_3 , pour $\mathcal{N}=9$. (d) Résidus de convergence moyens pour $\mathcal{N}=35$. (e) Images des zones d'intérêt centrées autour du point A : avant (première colonne) et après (deuxième colonne) correction par P_6 . (f) Paramètres d'illumination de P_3 . (g) Évolution des paramètres d'illumination de P_6 .

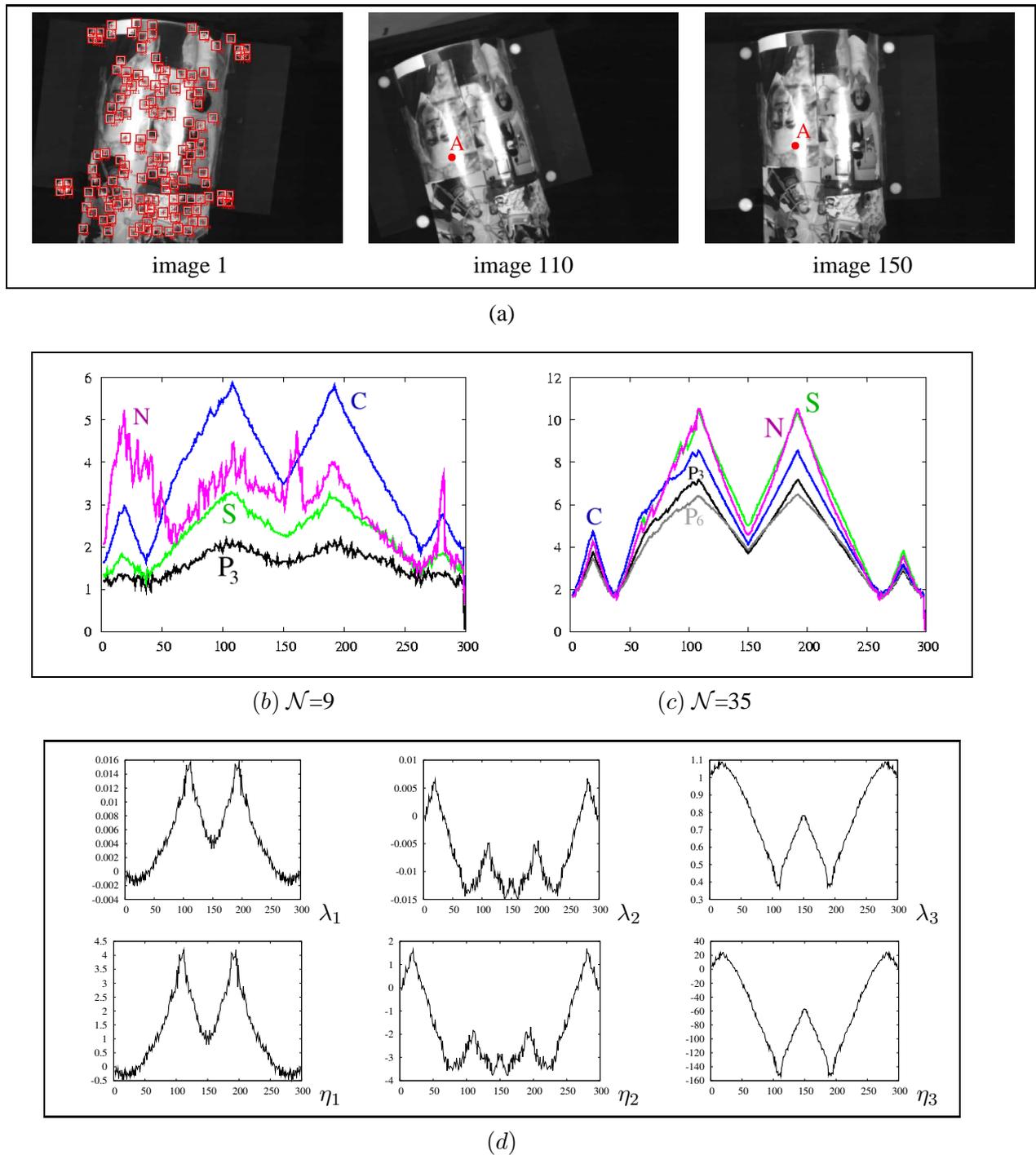


FIG. 7.9 – Séquence Cylindre. (a) Images de la séquence (137 points ont été sélectionnés). (b) Évolution de la moyenne des résidus pour $\mathcal{N} = 9$. (c) Évolution de la moyenne des résidus pour les points pour $\mathcal{N} = 35$. (d) Évolution des paramètres photométriques calculés au point A par P_6 .

TAB. 7.2 – Pourcentage de points suivis jusqu'à la fin de la séquence par rapport aux points sélectionnés, en fonction de la taille de fenêtre \mathcal{N} , dans le cas d'apparition de spécularités.

(a) Livre (97 points sélectionnés)							(b) Cylindre (137 points sélectionnés)						
\mathcal{N}	9	11	13	15	25	35	\mathcal{N}	9	11	13	15	25	35
C	27.3	30.7	23	18.6	11.8	9.4	C	86.5	83.5	81.1	80.2	71.4	60.5
N	37.5	53.4	60.9	65.1	63.5	50.6	N	40.6	54.9	62.1	69.5	83.3	71.4
S	17	28.4	55.2	65.1	77.6	81.2	S	76.7	80.5	81.8	85.5	85.7	73.9
P_3	77.3	77.3	79.3	77.9	78.8	76.5	P_3	96.2	94	93.9	93.1	87.3	77.3
P_6	-	-	34.5	54.7	90.6	83.5	P_6	-	-	51.5	64.8	81.7	88.2

Extension aux motifs très larges. La figure 7.10a représente une séquence d'images montrant un objet spéculaire de forme non plane. Une zone de l'image est sélectionnée² au voisinage de la région de forte saturation correspondant à une forte spécularité. Avec $\mathcal{E}_{moy} = 25$, seules les approches P_3 et P_6 permettent de suivre le motif du début à la fin de la séquence. En outre, la figure 7.10b, qui représente les résidus de convergence obtenus, montre que P_6 modélise plus précisément les variations de spécularité provoquées comparativement à P_3 , puisqu'elle aboutit à des résidus plus faibles. Notons également que C est l'approche qui perd le motif le plus rapidement en comparaison à N et S .

Les expérimentations présentées jusqu'à présent ont été effectuées sur des séquences d'images montrant des spécularités. La section suivante s'intéresse quant à elle à la comparaison des approches de suivi dans le cas où des variations d'éclairage sont également provoquées.

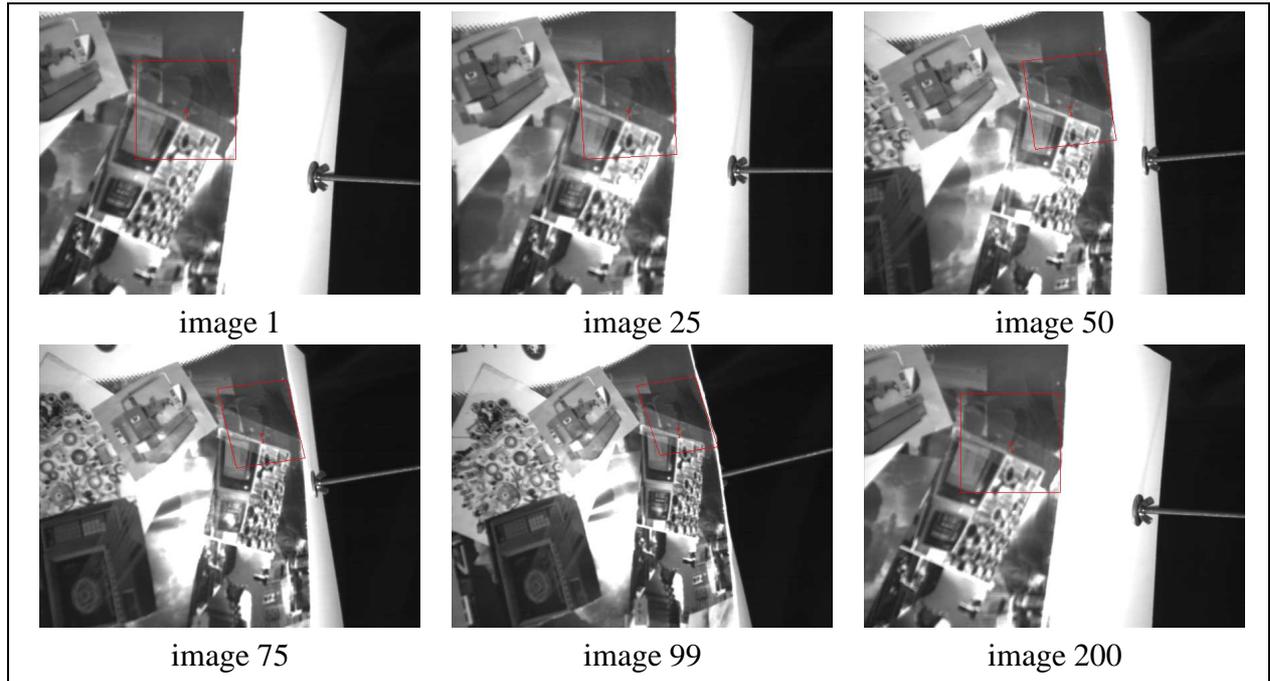
7.6.2.2 Variations d'éclairage et de spécularité

Nous considérons ici quatre séquences d'images, montrant des variations d'éclairage, et dans certains cas des spécularités. Les séquences *Objet Plan* et *Marylin* représentent des objets texturés (voir les figures 7.11a et 7.12a), constitués de différents matériaux (papier glacé, céramique, métal, carton, verre) éclairés par un éclairage ambiant (la lumière du jour et des lampes fluorescentes situées au plafond) et une source d'éclairage direct. Ensuite, la séquence *Hill* (figure 7.13a) et la séquence *Corner*³(figure 7.14a) montrent des scènes d'extérieur acquises *a priori* à différents moments de la journée. Dans chacun des cas, la caméra se déplace tandis que la scène est fixe.

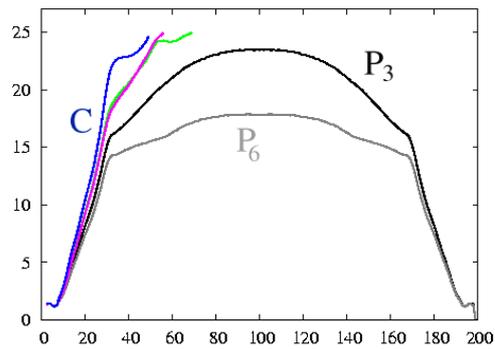
Dans la séquence *Objet Plan*, l'éclairage direct subit des variations d'intensité très brutales de façon périodique, à une période de 20 itérations environ, d'une valeur maximale à une valeur minimale. La séquence *Marylin* est particulièrement compliquée du fait du large déplacement de la caméra et des occultations engendrées. De plus, des variations d'intensité de la source d'éclairage sont là encore volontairement provoquées : autour de l'itération 135, la lumière directe est éteinte, induisant de brusques variations de l'illumination. Pour finir, les séquences

²Ici, nous sélectionnons « manuellement » le motif.

³Ces séquences font partie de la base d'images CMU/VASC : <http://vasc.ri.cmu.edu/idb/html/motion/-index.html>



(a)



(b)

FIG. 7.10 – Prise en compte de larges motifs ($\mathcal{N} = 151$), apparition de spéularités. (a) Images de la séquence considérée et motif suivi par P_6 . (b) Évolution des résidus de convergence.

TAB. 7.3 – Temps de calcul (en ms) du suivi d'un point dans la séquence Livre, pour $\mathcal{N}=9, 15$ et 35 .

Méthode	$\mathcal{N}=9$	$\mathcal{N}=15$	$\mathcal{N}=35$
C	1.3	2.7	21
N	4.6	6.8	31.2
S	1.7	3.1	21.7
P_3	1.4	3.2	12.3
P_6	9	8.5	25

Hill et *Corner* font apparaître des changements de lumière qui ne sont pas homogènes dans l'image. Dans un premier temps, nous considérons des fenêtres de taille inférieure à $\mathcal{N} = 35$. Là aussi, nous étudions la robustesse du suivi et la précision de la modélisation. Une dernière expérimentation sur la séquence *Objet Plan* va permettre d'analyser les performances du suivi de larges motifs.

Robustesse. On sélectionne 58 points dans la première image de la séquence *Objet Plan*, 156 dans *Hill*, 56 points dans *Marylin* et enfin 44 dans *Corner*. Ces valeurs sont regroupées dans les tableaux 7.4a, 7.4b, 7.4c, 7.4d. Pour de petites tailles de fenêtres d'intérêt ($\mathcal{N} < 15$), et quelle que soit la séquence d'image considérée, P_3 perd le moins de points. Par contre, à partir de $\mathcal{N} = 25$, la méthode P_6 s'avère à chaque fois la plus robuste. Globalement, pour de grandes fenêtres d'intérêt, chacune des approches, hormis la technique C , montre une bonne robustesse du suivi. En comparaison par rapport aux expériences précédentes où seules des variations de spécularités étaient provoquées, les méthodes N et S montrent ici une meilleure robustesse du suivi. En effet, elles s'avèrent dans tous les cas plus robustes que C , alors que ce n'était pas le cas pour de petites fenêtres d'intérêt lorsque seules des spécularités apparaissaient. Ces approches sont donc mieux adaptées pour compenser les variations d'éclairage que pour prendre en compte les occurrences de spécularités.

Précision des modèles. Les résidus de convergence obtenus pour $\mathcal{N} = 9$ par les différentes approches de suivi sont montrés sur les figures 7.11b, 7.12b, 7.13b et 7.14b. Ces résidus évoluent globalement de la même manière que les variations d'illumination provoquées. Cela se distingue très nettement sur la figure 7.11b, où ils varient de manière périodique, à la même fréquence que les changements d'intensité provoqués (20 itérations environ entre les valeurs maximales et minimales). Dans la séquence *Marylin*, seule P_3 a suivi des points (figure 7.12b). Pour les autres séquences, les moyennes des résidus de S , N et P_3 sont très comparables. Or ces moyennes de résidus sont calculées sur P_3 sur un plus grand nombre de points que S et N (voir les tableaux 7.4a, 7.4c et 7.4d). Pour affiner la comparaison, les figures 7.11c et 7.13c comparent les résidus obtenus par P_3 , N et S sur les quelques points suivis correctement par l'ensemble de ces trois approches. Pour la séquence *Objet plan*, P_3 obtient des résidus plus faibles que N . Dans le cas de la séquence *Hill*, il est plus difficile d'aboutir à une conclusion, dans le sens où les résidus sont très comparables. Dans la séquence *Marylin*, pour $\mathcal{N} = 15$ (voir la figure 7.12c), les résidus de P_3 sont plus élevés que ceux des autres techniques. P_3 est donc plus appropriée sur de petites fenêtres, dans le cas de changements d'éclairage. Pour des fenêtres plus larges, quelle que soit la séquence d'images considérée (voir l'évolution des résidus sur les figures 7.11d, 7.12d, 7.13d et 7.14c) la méthode P_6 détient les résidus de convergence les plus faibles. Ces résultats montrent finalement la bonne modélisation des changements d'illumination par P_3 pour de petites fenêtres et par P_6 dans le cas de plus grands motifs. En effet, cette approche permet de compenser les variations spatiales des changements d'illumination. Pour les séquences *Objet plan*, *Marylin* et *Corner*, P_3 obtient de moins bons résidus que S , N et P_6 . Par contre, sur la séquence *Hill*, elle aboutit à de plus faibles résidus. Les trois premières séquences correspondent à des scènes fortement structurées, où la reflectance est susceptible de varier localement de manière discontinue (forts contours) tandis que la séquence *Hill* s'avère plus texturée et montre peu de fortes discontinuités de la reflectance. Dans ce cas de figure, P_3 aboutit à des résidus de convergence plus faibles que N et S . Par conséquent, il est légitime de penser que, dans le cas où des variations d'éclairage sont provoquées, les performances

de P_3 dépendent des variations de réflectance de la surface considérée. Du fait de l'hypothèse formulée quant à la continuité des variations d'illumination (voir la section 6.3.1), plus la réflectance montre de fortes discontinuités et moins la modélisation proposée parvient à compenser ces variations. Dans le cas des expérimentations précédentes, où seules des spécularités étaient considérées, P_3 détenait globalement des résidus plus faibles que N et S quelle que soit la taille de fenêtre (voir les figures 7.8d et 7.9c par exemple). Ainsi, P_3 s'avère plus pertinente pour prendre en compte les spécularités que pour compenser les variations d'intensité d'éclairage, où le formalisme utilisé doit compenser des variations dépendant de la réflectance.

Paramètres photométriques. La première rangée d'images de la figure 7.11e montre les intensités de la fenêtre d'intérêt centrée autour d'un des points suivis, sélectionné dans la zone de fortes variations d'illumination (point A visible sur la première image de la figure 7.11a). Les variations brutales d'éclairage sont bien remarquables. La deuxième rangée d'images est associée aux intensités corrigées par la modélisation photométrique utilisée par l'approche P_6 . Les variations d'illumination ne sont plus perceptibles. Les paramètres associés, ηU^T et λU^T avec $U^T = (x - x_p, y - y_p, 1)$ sont représentés sur la figure 7.11f, respectivement sur la première et la seconde rangée. Ils permettent de compenser les variations spatiales de contraste et de variations de réflexion spéculaire. L'évolution de ces paramètres au cours de la séquence est montrée sur la figure 7.11g. Elle correspond effectivement à la variation d'éclairage provoquée. Nous retrouvons notamment une fréquence approximative de 20 itérations entre une valeur maximale d'intensité et une valeur minimale.

Temps de calcul. D'après les temps de calcul regroupés dans le tableau 7.5, les approches N et P_6 sont là encore les approches les plus coûteuses en temps de calcul. Pour $\mathcal{N} = 9$ et $\mathcal{N} = 15$, les temps de calcul de P_6 sont importants, car cette approche ne converge pas bien sur les petites fenêtres d'intérêt. Notons également que P_3 obtient des temps plus importants que dans le cas de spécularités (voir le tableau 7.3). Ceci tend à montrer que, même si elle permet une bonne robustesse du suivi, elle s'avère plus appropriée dans le cas de variations de spécularités que dans le cas de changements d'éclairage.

Intéressons-nous à présent au suivi de larges zones de l'image.

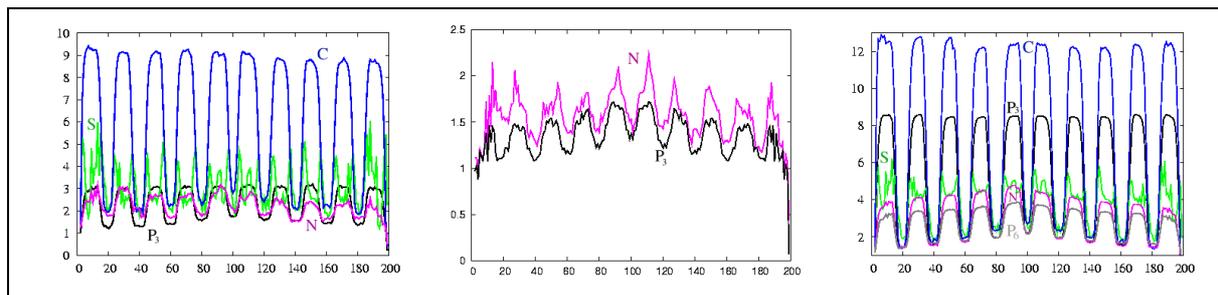
Extension au suivi de larges zones de l'image. La figure 7.15a représente une séquence d'image montrant l'objet plan de la séquence *Objet plan*. Une zone de l'image est sélectionnée, et le suivi est effectué avec $E_{moy} = 15$. La figure 7.15b représente les résidus de convergence obtenus. Ils montrent que P_6 modélise plus précisément les variations de spécularités provoquées que les autres approches, puisqu'elle aboutit à des résidus plus faibles. Nous n'avons pas représenté les résidus obtenus par C , puisque cette approche n'a pas permis de suivre le motif tout au long de la séquence. L'approche P_6 s'avère donc être la plus adaptée pour compenser les variations d'illumination sur de larges zones de l'image. Notons toutefois que les méthodes N , C et S n'ont pas été développées dans le but de prendre en compte des motifs aussi larges.

7.6.2.3 Bilan des expérimentations et remarques

Les résultats d'expérimentation obtenus confirment que la technique classique de suivi (C) n'est ni robuste aux variations de spécularités ni aux changements d'illumination, puisqu'elle se base sur l'hypothèse de conservation de la luminance. Globalement, l'utilisation d'un modèle



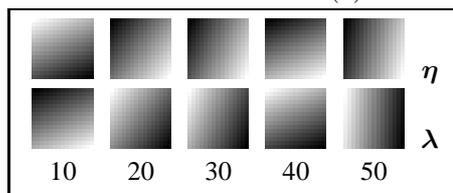
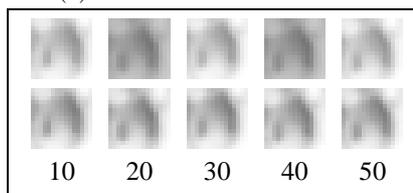
(a)



(b) $\mathcal{N} = 9$

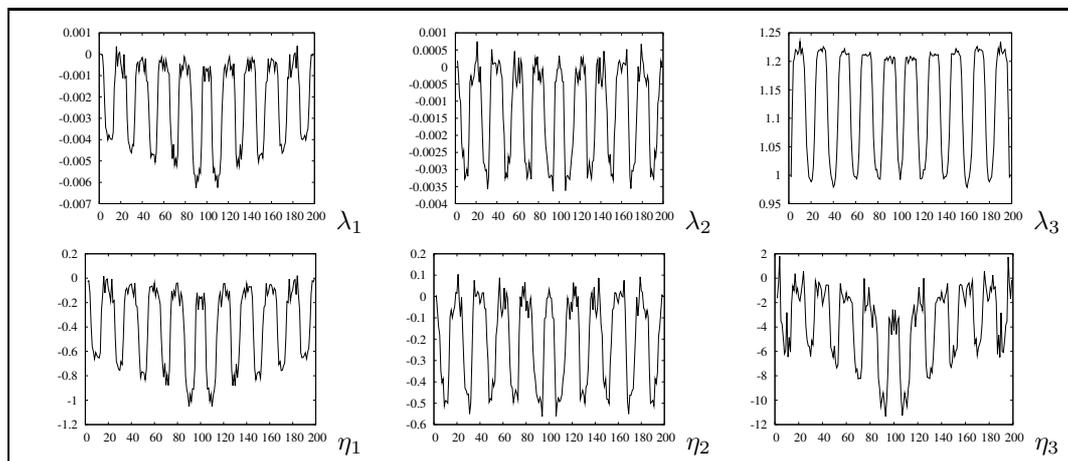
(c) $\mathcal{N} = 9$

(d) $\mathcal{N} = 35$



(e)

(g)



(f)

FIG. 7.11 – Expérimentation Objet plan. (a) Trois images de la séquence. (b) Résidus moyens pour $\mathcal{N}=9$. (c) Résidus obtenus pour $\mathcal{N}=9$ sur les points suivis pas N et P_3 . (d) Résidus moyens obtenus avec $\mathcal{N}=35$. (e) Images de la fenêtre d'intérêt centrée sur A : avant (première colonne) et après (deuxième colonne) correction par le modèle à 6 paramètres P_6 . (f) Paramètres d'illumination calculés avec P_6 sur la fenêtre d'intérêt. (g) Évolution des paramètres d'illumination calculés avec P_6 .

TAB. 7.4 – Pourcentage de points suivis au cours de la séquence, dans le cas de spécularités et de variations d'éclairage (les points occultés ou sortant de l'image ne sont pas pris en compte).

(a) <i>Objet Plan</i> (58 points sélectionnés)							(b) <i>Hill</i> (156 points sélectionnés)						
\mathcal{N}	9	11	13	15	25	35	\mathcal{N}	9	11	13	15	25	35
C	63.8	50	39.7	36.2	8.6	6.9	C	49.6	31.1	23.7	20	11.2	9.6
N	77.6	87.9	91.4	91.4	96.6	93.1	N	45.9	60	63.7	63.7	73.7	75
S	67.2	82.8	87.9	51	87.9	89.7	S	55.6	56.3	63.7	70.4	85.6	93.3
P_3	100	100	100	100	96.6	96.6	P_3	74.8	74.8	74.8	75.6	86.4	95.2
P_6	48.3	75.9	87.9	94.8	100	100	P_6	-	67.4	70.4	77.8	89	97.1

(c) <i>Marylin</i> (56 points sélectionnés)							(d) <i>Corner</i> (44 points sélectionnés)						
\mathcal{N}	9	11	13	15	25	29	\mathcal{N}	9	11	13	15	25	29
C	0	0	0	0	0	0	C	90.9	86.4	88.6	88.6	84.1	67.4
N	0	2.8	2.8	16.7	13.9	13.9	N	72.7	63.6	90.9	95.5	86.4	81.8
S	0	2.8	2.8	5.5	8.3	13.9	S	100	100	100	100	97.7	88.6
P_3	36.1	22.2	16.7	11	5.5	2.8	P_3	100	100	100	100	100	100
P_6	-	-	-	11.1	33.3	30.6	P_6	34.1	50	72.7	86.4	100	100

TAB. 7.5 – Temps de calcul (en ms) du suivi d'un point dans la séquence *Objet Plan*, pour $\mathcal{N}=9, 15$ et 35 .

Méthode	$\mathcal{N}=9$	$\mathcal{N}=15$	$\mathcal{N}=35$
C	1.3	2.9	11.4
N	4.3	3.5	14.1
S	1.6	3.2	11.5
P_3	2	3.5	13.8
P_6	32	5.9	18.4

photométrique affine (S et N) apporte une meilleure robustesse, excepté dans le cas d'occurrence de spécularités lorsque la fenêtre considérée est petite. Ce constat s'explique certes par une compensation moins adaptée des variations d'illumination, mais il faut également souligner leur sensibilité au bruit. En effet, dans le cas d'un pixel bruité dans \mathcal{W} , les valeurs μ_f , σ_f , μ_g , σ_g , ainsi que λ deviennent également bruitées, puisqu'elles dépendent de chacune des luminances de \mathcal{W} . Pour l'approche S , λ est multipliée par chaque valeur de f , et une erreur sur λ a par conséquent une grande influence. La minimisation de ϵ_2 et ϵ_3 peut mener finalement à une valeur incorrecte de \mathbf{A} . En revanche, pour de larges fenêtres d'intérêt, la contribution d'un pixel bruité dans le calcul de ces paramètres devient moins significative, le calcul de μ_f , μ_g , σ_f , σ_g , λ est plus précis et par conséquent \mathbf{A} est également plus précis. La section C.1 de l'annexe C montre à cet effet l'erreur provoquée par un pixel bruité sur le calcul de la normalisation photométrique effectuée par N .

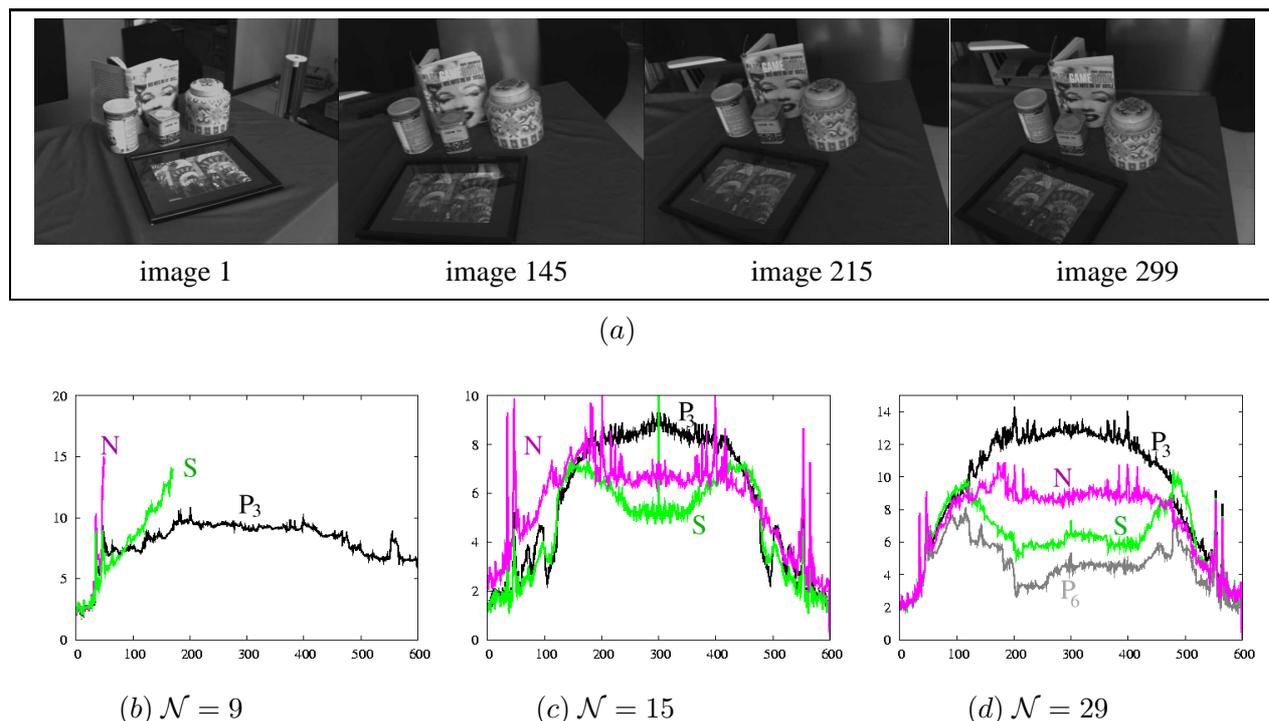


FIG. 7.12 – Expérimentation Marilyn. (a) Images de la séquence. (b) Résidus moyens obtenus pour $\mathcal{N}=9$. (c) Résidus moyens pour $\mathcal{N}=15$. (d) Résidus moyens pour $\mathcal{N}=29$.

Par ailleurs, l'approche S peut conduire à une matrice à inverser mal conditionnée pour certains points, étant donné que la matrice mise en jeu dépend des luminances (voir la section 7.3.1). La section C.2 de l'annexe C prouve d'ailleurs que cette technique obtient un plus mauvais conditionnement que l'approche P_3 , dans le cas d'un motif formant un coin idéal, primitive pourtant facilement détectée par la méthode de Harris. En guise d'exemple, nous avons comparé le conditionnement en quelques points (10 au total) sélectionnés dans la séquence *Cylindre*. Le tableau 7.6 regroupe les rapports :

Conditionnement de la méthode / Conditionnement de la méthode classique,

obtenus sur des tailles de fenêtre $\mathcal{N}=9, 15$ et 35 . Plus ces valeurs sont élevées et moins la matrice mise en jeu est correctement conditionnée. En outre, nous précisons les valeurs de conditionnement maximale (Max), minimale (Min) et moyenne (Moy) obtenues sur les 10 points. Notons que le conditionnement de N est identique à l'approche classique de suivi. D'après les résultats du tableau 7.6, les techniques S , P_3 et P_6 sont toutes moins bien conditionnées que l'approche classique. En particulier pour de petites fenêtres ($\mathcal{N} = 9$), les techniques S et P_6 sont très mal conditionnées, comparativement aux approches classiques et à l'approche P_3 . Cette dernière obtient toutefois un meilleur conditionnement que P_6 et S .

Pour de petites fenêtres d'intérêt, la méthode P_3 suit un plus grand nombre de points que N et S . Dans ce cas de figure, elle compense convenablement les variations de spéularité et

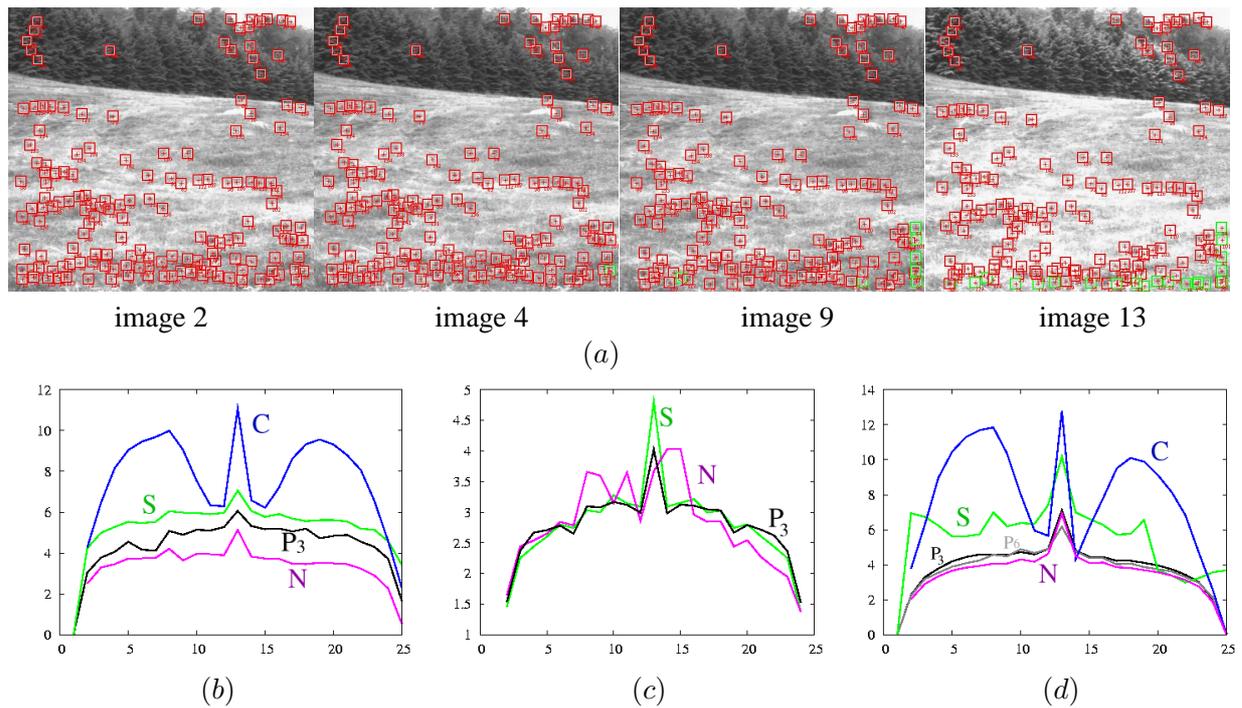


FIG. 7.13 – Expérimentation Hill. (a) Images de la séquence. (b) Résidus moyens pour $\mathcal{N} = 9$. (c) : résidus pour $\mathcal{N} = 15$ obtenus sur les points suivis à la fois par P_3 , N et S . (d) Résidus moyens pour $\mathcal{N} = 35$.

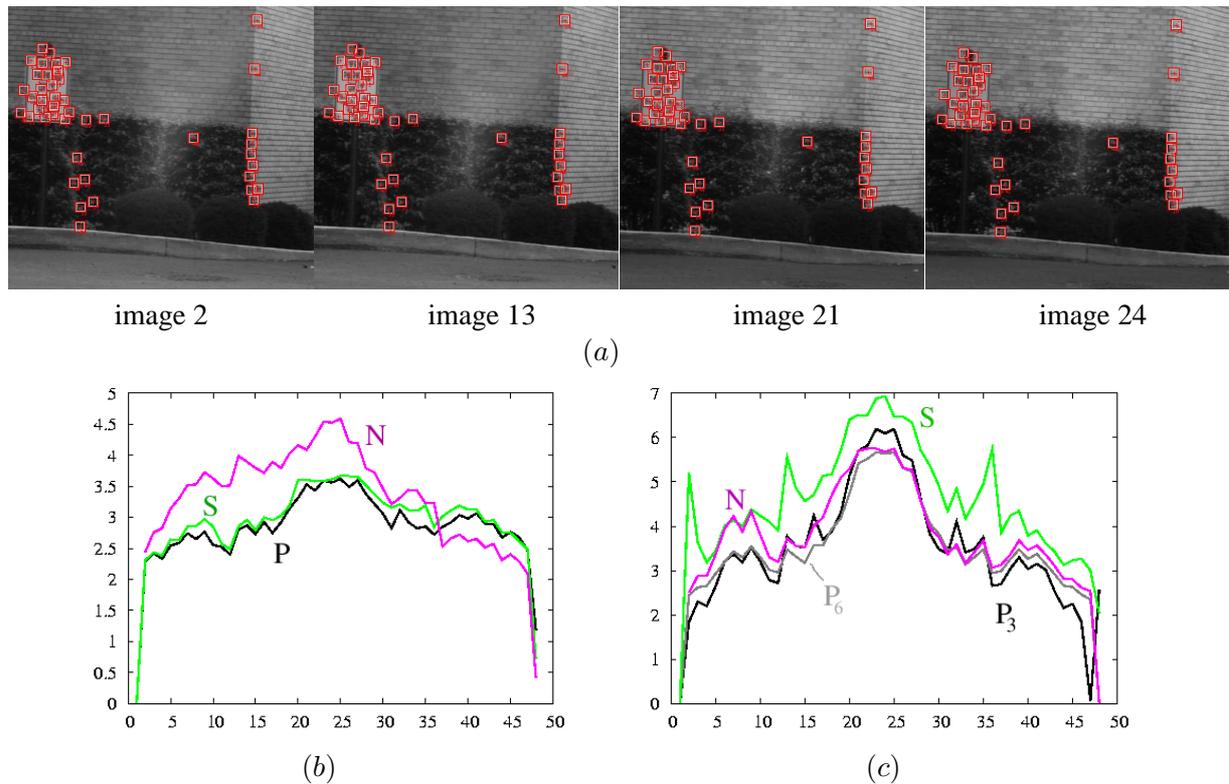


FIG. 7.14 – Expérimentation Corner. (a) Quelques images de la séquence. (b) Résidus moyens pour $\mathcal{N} = 9$. (c) Résidus moyens pour $\mathcal{N} = 35$.

TAB. 7.6 – Conditionnement des matrices : rapport du nombre de conditionnement de la méthode sur celui de la méthode classique, calculé sur 10 points de la séquence *Cylindre*.

	$\mathcal{N}=9$			$\mathcal{N}=15$			$\mathcal{N}=35$		
	Max	Min	Moy	Max	Min	Moy	Max	Min	Moy
C	1	1	1	1	1	1	1	1	1
S	9	1	3.5	3.6	1	1.4	1.5	1.1	1.1
P_3	2	0.5	1.4	3.1	0.6	1.2	1.4	0.5	0.8
P_6	-	-	-	4.2	1.2	2.3	1.7	1.1	1.3

Max : Valeur de conditionnement maximale (parmi les 10 points considérés)

Min : Valeur de conditionnement minimale

Moy : Valeur de conditionnement moyen

d'éclairage sur \mathcal{W} . Par contre, ses performances se dégradent lorsqu'il s'agit de compenser des changements d'éclairage sur de grandes tailles de fenêtre. En effet, dans ce cas de figure, la modélisation prise en compte doit approximer une partie de la reflectance diffuse de l'objet par un polynôme de degré 1 sur \mathcal{W} . Or cette hypothèse peut s'avérer grossière sur de larges fenêtres où la reflectance est susceptible de montrer de fortes variations.

Notons également que, du point de vue des temps de calcul, même si P_3 requiert le calcul d'un paramètre supplémentaire par rapport à S et donc l'inversion d'une matrice plus grande, les temps de ces deux approches sont similaires, du fait de la meilleure convergence de P_3 .

Enfin, l'approche P_6 s'avère appropriée pour de larges fenêtres, quelles que soient les variations d'illumination. Toutefois, son utilisation sur des tailles de fenêtre restreintes ne permet pas de calculer précisément les six paramètres photométriques mis en jeu.

En définitive, ces résultats d'expérimentation corroborent les résultats de simulation obtenus dans la section 7.6.1.

7.7 Conclusions

Puisque l'utilisation de modèles de réflexion spéculaires induit la prise en compte d'un grand nombre de paramètres, la plupart des algorithmes de vision par ordinateur suppose que les objets sont lambertiens et qu'aucun changement de spécularités ou d'éclairage ne se produit. Il s'agit bien sûr d'une hypothèse non réaliste.

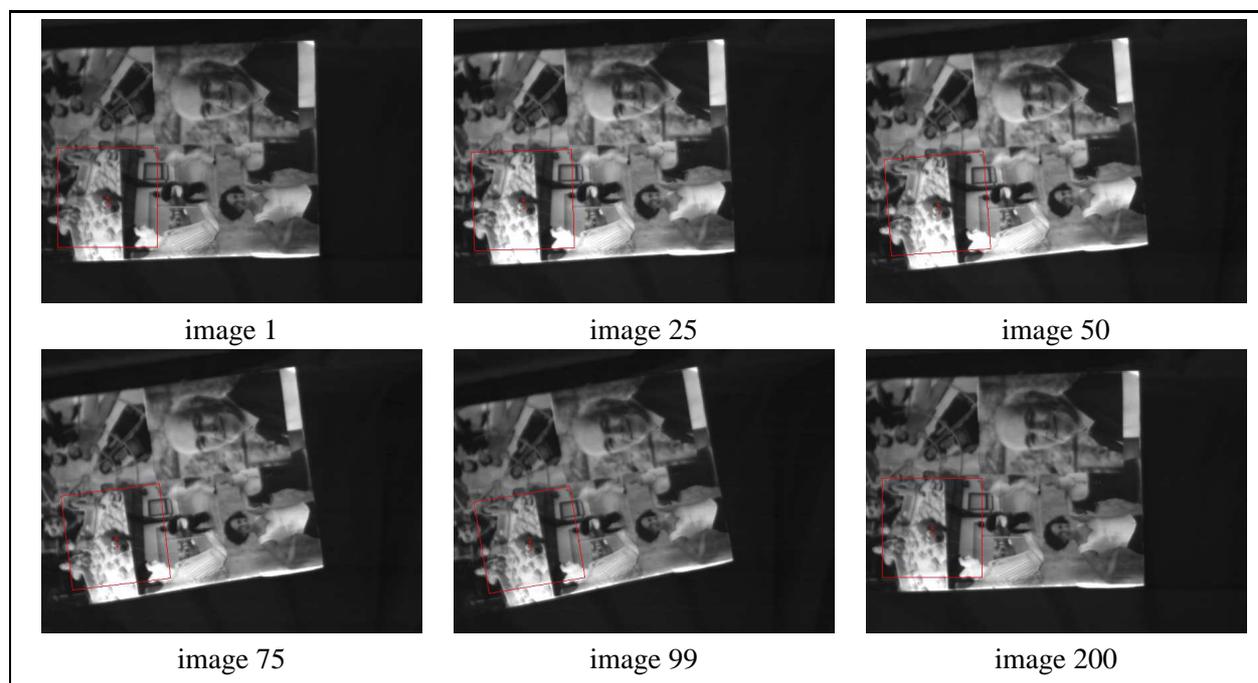
Néanmoins, l'utilisation de modèles photométriques simplifiés sur des fenêtres d'intérêt de taille réduite améliore de manière significative la robustesse des traitements, en prenant en compte les variations d'illumination entre les images de la séquence. Dans ce chapitre, nous avons exploité les deux modèles d'illumination proposés dans le chapitre 6 dans le cadre du suivi de points d'intérêt et du suivi de motifs plus larges. Ces approches ont été comparées, d'un point de vue de la robustesse et de la précision, aux approches de suivi de points classiquement utilisées dans la littérature : la technique classique KLT, les méthodes compensant des

variations d'illumination affine. En prenant en compte les variations spatiales de changements d'illumination, les approches proposées s'avèrent plus robustes que ces techniques existantes.

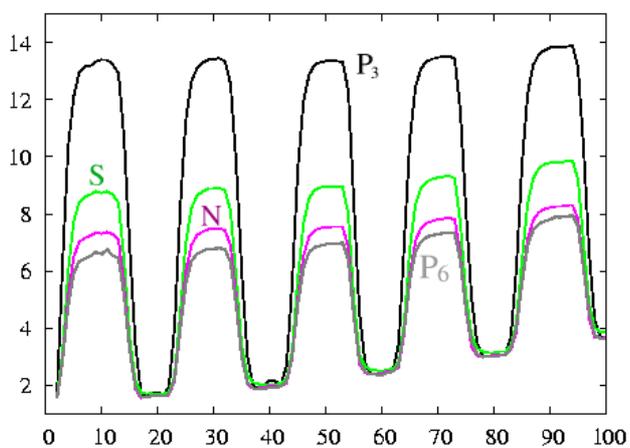
La première technique proposée permet d'améliorer la robustesse du suivi de points vis-à-vis des variations de spécularités et des changements d'éclairage. Dans le cadre du type d'applications auxquelles nous nous intéressons au Cemagref, cette approche pourra permettre d'améliorer la robustesse de techniques d'asservissement visuel, comme dans [Collewet et Chaumette, 2002], où un positionnement de la caméra en face d'une zone d'intérêt est effectué, à l'aide de trois points d'intérêt.

La seconde approche s'applique à de larges fenêtres d'intérêt. Elle peut par exemple être exploitée dans le cadre du calcul de la structure de l'objet à partir du mouvement. Dans cette optique, les travaux décrits dans [Alhaj, 2004] se basent sur une zone de taille 101×101 permettant de calculer les paramètres de mouvement et la forme de l'objet (localement) afin d'assurer le positionnement du capteur parallèlement au plan tangent en un point particulier de l'objet.

Après avoir traité le problème du suivi de points dans les images de luminance, intéressons-nous à présent au cas des images couleur. En effet, l'information colorimétrique permet, moyennant quelques hypothèses, de déduire des attributs invariants valables globalement en chaque pixel de l'image. Le chapitre suivant propose un état de l'art de ces attributs. Certains d'entre eux seront par la suite exploités dans le cadre du suivi de points d'intérêt.



(a)



(b)

FIG. 7.15 – Changements d'éclairage et de spécularités, prise en compte de larges motifs ($\mathcal{N} = 151$). (a) Images de la séquence considérée et motif suivi par P_6 . (b) Évolution des résidus de convergence.

L'INVARIANCE COULEUR

Contrairement aux images de luminance, pour lesquelles la compensation des variations d'illumination est obtenue par le biais de modélisations valables localement dans l'image (voir le chapitre 6), l'utilisation de la couleur permet de définir des attributs invariants vis-à-vis de ces phénomènes de manière globale, en chaque point de l'image. Ce chapitre décrit les principaux invariants couleur existants, en fonction des caractéristiques optiques des matériaux. Ces attributs seront utilisés dans le chapitre 9, dans le cadre du suivi de points dans les images couleur.



La couleur perçue d'un objet n'est jamais parfaitement uniforme et dépend énormément, comme nous l'avons vu dans le chapitre 1, des conditions d'éclairage (la couleur de l'illuminant, son intensité, sa direction) et d'observation. L'œil humain parvient très bien à percevoir l'uniformité de la couleur quelles que soient la direction d'éclairage, l'intensité de la source et même sa couleur, mais ce problème d'invariance reste posé en vision par ordinateur. Ici, nous nous intéressons principalement à l'invariance des couleurs d'un objet par rapport à la géométrie de la scène, c'est-à-dire par rapport aux positions relatives capteur-éclairage-objet, ainsi qu'aux changements d'intensité de la source. Quant à la *constance couleur*, c'est-à-dire l'invariance vis-à-vis des changements de couleur de l'illuminant, qui nécessite généralement la détermination de cette couleur, elle n'est pas abordée ici. On pourra se référer par exemple aux travaux de Barnard *et al.* [Barnard *et al.*, 1997], Finlayson [Finlayson *et al.*, 2000] et aux travaux effectués dans [Brill, 1990, Geusebroek *et al.*, 2003], où ce problème est traité.

L'invariance couleur a suscité beaucoup d'intérêt et continue à intéresser le domaine de la vision par ordinateur, et tout particulièrement les applications de reconnaissance d'objets. En effet, c'est certainement dans ce domaine que l'utilisation des invariants couleur s'avère le plus intéressant. Il s'agit alors d'être capable de retrouver correctement l'image d'un objet dans une base d'images, bien que les conditions d'acquisition soient différentes [Ohba et Ikeuchi, 2000, Gevers et Smeulders, 1997a, Drew *et al.*, 1998, R. Alferez, 1999, Kim *et al.*, 2001, Gevers, 2001, Lin *et al.*, 2002]. Le plus souvent, des hypothèses sont effectuées sur le type d'objet de

manière à utiliser des invariants couleur adaptés [Gevers et Smeulders, 1997a, Ohba et Ikeuchi, 2000]. Ces attributs sont également exploitables en indexation d'images, comme dans les travaux présentés dans [Gevers et Smeulders, 2000] où invariants couleur et invariants de forme sont combinés. La segmentation d'image peut également être largement améliorée par l'utilisation d'attributs invariants en assurant une segmentation qui soit fidèle aux objets. En utilisant des composantes couleur classiques, les zones d'ombres ou les zones de spécularités peuvent être assimilées à des régions à part entière bien qu'elles ne correspondent à aucun objet de la scène. Citons notamment les travaux précurseurs de Healey [Healey, 1992] ou les travaux plus récents de [Gevers, 2002, Gevers et Smeulders, 1997b, Geusebroek *et al.*, 2000a, Gevers *et al.*, 1998, Drew *et al.*, 1999]. Dans [Gevers et Stokman, 2003, Gevers, 2003a] est proposée une méthode permettant de classifier les contours par type : contours de spécularités, contours d'ombres, contours d'objets. Les invariants couleur ont également été utilisés dans des tâches de mise en correspondance d'images. Dans [Gevers et Smeulders, 1996] on peut trouver une comparaison de quelques attributs invariants. Certains travaux utilisent des invariants couleur dans le cadre d'une estimation de mouvement dans une vidéo [Golland et Bruckstein, 1997]. Dans cet article, les auteurs prouvent la pertinence et la meilleure précision des attributs couleur invariants pour le calcul de mouvement. La robustesse des traitements est en effet fortement améliorée par l'invariance vis-à-vis des changements d'intensité d'éclairage ou de couleur d'illuminant survenant au cours de la séquence d'images.

La plupart des méthodes de calcul d'invariants se base sur le modèle dichromatique de Shafer (voir la section 1.2.2.6 page 23), en effectuant des hypothèses quant aux variables impliquées dans ce modèle. Nous abordons ici les invariants dédiés aux matériaux diélectriques inhomogènes opaques, généralement dits *lambertiens*, puis aux invariants valables pour les matériaux conducteurs opaques dits *spéculaires*. Dans le seul but d'alléger les relations exprimées dans ce chapitre, nous considérons que l'irradiance reçue par le capteur correspond directement à la radiance de l'objet. Gardons toutefois en mémoire que théoriquement ces deux grandeurs sont reliées par une constante (voir la relation (2.1)).

8.1 Modélisation de la couleur

Le modèle dichromatique exprime la radiance comme la somme de deux radiances :

$$\mathcal{L}(\lambda, P) = l_s(\lambda, P)m_s(P) + l_b(\lambda, P)m_b(P) \quad (8.1)$$

où l_b et l_s correspondent aux radiances spectrales respectivement de réflexion diffuse et de réflexion spéculaire. $m_b(p)$ et $m_s(p)$ expriment quant à eux les termes de géométrie de la scène. Rappelons que le modèle dichromatique permet d'écrire la couleur d'un matériau de la manière suivante :

$$\mathbf{C}(p) = m_b(p)\mathbf{C}_b(p) + m_s(p)\mathbf{C}_s(p) \quad (8.2)$$

où C_b et C_s correspondent respectivement à la couleur associée à la réflexion diffuse (*body reflection* en anglais) et la couleur associée à la réflexion spéculaire (voir le chapitre 1). Rappelons également que les couleurs C_b et C_s sont données par :

$$C_b^i(p) = \int_{\lambda} S_i(\lambda) \mathcal{E}(\lambda, P) \mathcal{R}_b(\lambda, P) d\lambda, \quad C_s^i(\lambda, p) = \int_{\lambda} S_i(\lambda) \mathcal{E}(\lambda, P) d\lambda \quad (8.3)$$

où $\mathcal{R}_b(\lambda, P)$ désigne la reflectance diffuse du matériau.

Dans [Geusebroek *et al.*, 2001, Geusebroek *et al.*, 2000b], le modèle de radiance résultant de la théorie de Kubelka-Munk est défini de la manière suivante au point P :

$$\mathcal{L}(\lambda, P) = \mathcal{E}(\lambda, P)(1 - \rho_f(P))^2 \mathcal{R}_b(\lambda, P) + \mathcal{E}(\lambda, P) \rho_f(P) \quad (8.4)$$

où $\rho_f(P)$ est la reflectance de Fresnel au point P . L'équation (8.4) équivaut au modèle de Shafer [Shafer, 1985], en posant $l_b(\lambda, P) = \mathcal{E}(\lambda, P) \mathcal{R}_b(\lambda, P)$ et $l_s(\lambda, P) = \mathcal{E}(\lambda, P)$ ainsi que $m_b(P) = (1 - \rho_f(P))^2$ et $m_s(P) = \rho_f(P)$. La plupart des travaux sur les invariants couleur se placent dans l'hypothèse selon laquelle le modèle dichromatique est respecté. Nous verrons que les différents raisonnements découlent soit de la relation (8.2), soit de la relation (8.4).

Ainsi, en fonction des hypothèses effectuées sur les conditions d'acquisition (illuminant d'énergie uniforme ou non), sur le type de matériau considéré (matériau mate ou non), sur sa couleur (uniforme ou non), différents invariants peuvent être formulés. On différenciera notamment le cas d'un illuminant blanc de celui d'un illuminant coloré. Certes, il n'existe pas d'illuminant parfaitement blanc mais, en admettant que l'intensité au point p , ne dépend pas des longueurs d'ondes $\mathcal{E}(\lambda, p) = I(p)$, cette hypothèse aboutit à l'expression d'invariants très utiles. Les composantes couleur du modèle de Shafer sont simplifiées pour $i \in \{R, G, B\}$ de la façon suivante :

$$C_b^i(\lambda, p) = I(p) \int_{\lambda} S_i(\lambda) \mathcal{R}_b(\lambda, P) d\lambda, \quad C_s^i(\lambda, p) = I(p) \int_{\lambda} S_i(\lambda) d\lambda \quad (8.5)$$

Dans un souci de lisibilité, appelons $D_i(\lambda)$ l'intégrale $\int_{\lambda} S_i(\lambda) \mathcal{R}_b(\lambda, P) d\lambda$, de sorte que chaque composante couleur du matériau s'écrit $C_b^i(\lambda, p) = I(p) D_i(\lambda)$. Qui plus est, l'intégrale des sensibilités spectrales des capteurs est identique pour chaque capteur lorsqu'une *balance des blancs** a été préalablement effectuée [Gevers et Smeulders, 1997a] :

$$\int_{\lambda} S_R(\lambda) d\lambda = \int_{\lambda} S_G(\lambda) d\lambda = \int_{\lambda} S_B(\lambda) d\lambda = Q \quad (8.6)$$

Dans ces conditions, le modèle dichromatique (8.2) devient :

$$\mathbf{C}(p) = I(p) m_b(p) \mathbf{D}(p) + I(p) m_s(p) Q \quad (8.7)$$

avec $\mathbf{D} = (D_R, D_G, D_B)^T$. Suivant les hypothèses effectuées sur le matériau observé, de nouvelles simplifications sont obtenues quant à la formation de l'image. Dans un premier temps, nous nous intéressons aux matériaux diélectriques, inhomogènes et opaques puis nous abordons le cas des objets conducteurs, dit spéculaires.

8.2 Invariants dédiés aux matériaux diélectriques, inhomogènes et opaques

Les matériaux diélectriques, inhomogènes et opaques (matériaux dits lambertiens) possèdent un coefficient de Fresnel ρ_f proche de 0. Ainsi la relation (8.4) devient :

$$\mathcal{L}(\lambda, P) = \mathcal{E}(\lambda, P)\mathcal{R}_b(\lambda, P) \quad (8.8)$$

Dans [Golland et Bruckstein, 1997], cette expression est appelée *loi de conservation de la couleur*. De même, le modèle dichromatique de l'équation (8.2) est très fortement simplifié dans le cas d'un objet lambertien, étant donné que la composante spéculaire est éliminée de sorte que $C(p) = m_b(p)C_b(p)$. La section suivante montre qu'en effectuant l'hypothèse selon laquelle l'illuminant est blanc, plusieurs invariants photométriques peuvent aisément être définis.

8.2.1 Illuminant blanc

En considérant un objet diélectrique inhomogène opaque et un illuminant blanc, tel que dans (8.7), la modélisation se simplifie encore puisque :

$$C(p) = I(p)m_b(p)D(p) \quad (8.9)$$

Cette simplification permet de définir des attributs invariants, qui sont détaillés ci-dessous.

Les rapports de composantes RGB. Posons tout d'abord $K(p) = m_b(p)I(p)$. D'après (8.9), le vecteur couleur C s'exprime comme le produit d'une composante géométrique $K(p)$ dépendant du positionnement relatif entre l'éclairage, la surface et la caméra et d'une composante spectrale D dépendant à la fois des sensibilités spectrales du capteur et du spectre de reflectance du matériau. Or, la sensibilité du capteur est une propriété intrinsèque, qui ne varie pas brutalement au cours du temps. D'autre part, le spectre de la lumière incidente est considéré d'allure constante, c'est-à-dire qu'il reste blanc au cours du temps. Pour finir, les caractéristiques spectrales de l'objet sont fixes¹. Ainsi les composantes D_i satisfont l'hypothèse d'invariance : ils sont invariants aux déplacements de l'objet, de la caméra ou des sources d'éclairage. De même, tout rapport de deux composantes couleur R , G et B satisfait l'hypothèse d'invariance, par élimination de la composante $K(p)$ [Healey, 1989]. Par exemple, les composantes normalisées par la norme L_1 (les coordonnées *rgb*, décrites par (2.10) page 34) sont invariantes. Puisque toute combinaison d'invariants est un invariant [Gevers et Smeulders, 2000], toute composante C_{inv} pouvant s'exprimer sous la forme :

$$C_{inv} = \frac{\sum_i a_i (R)_i^q (G)_i^r (B)_i^s}{\sum_j b_j (R)_j^t (G)_j^u (B)_j^v} \quad (8.10)$$

est un invariant par rapport à l'orientation de surface, la direction d'observation, la direction de l'illumination et l'intensité d'éclairage, dans le cas où l'objet est lambertien et l'illuminant

¹Si l'on ne considère pas des objets particuliers tels que les caméléons !

blanc. Dans (8.10), $(R)_i^q$ désigne le i -ième terme de la combinaison linéaire, et q la puissance à laquelle il est élevé. Enfin, $q + r + s = t + u + v$, pour $q, r, s, t, u, v, a_i, b_i \in \mathbb{R}, i, j \geq 1$. Soulignons que les rapports de composantes couleur RGB sont très sensibles au bruit dans le cas où l'intensité des composantes du dénominateur est très faible.

La norme L_2 . La normalisation L_2 des composantes R, G, B permet également d'aboutir à des invariants photométriques. Soit une composante couleur $C_i \in \{R, G, B\}$. La normalisation L_2 , $C_i^{L_2}$ de chaque composante C_i est donnée par :

$$C_i^{L_2} = \frac{C_i}{\sqrt{R^2 + G^2 + B^2}} = \frac{K(p)D_i(p)}{\sqrt{K(p)^2(D_R(p)^2 + D_G(p)^2 + D_B(p)^2)}} \quad (8.11)$$

$$= \frac{D_i(p)}{\sqrt{(D_R(p)^2 + D_G(p)^2 + D_B(p)^2)}} \quad (8.12)$$

Les composantes couleur RGB normalisées par la norme L_2 dépendent uniquement des sensibilités spectrales des capteurs et des caractéristiques physiques des matériaux, mais plus de la géométrie de la scène ni de l'intensité d'éclairage.

Les attributs $c_1 c_2 c_3$. Dans [Gevers et Smeulders, 1997a], l'auteur se base sur les propriétés du modèle dichromatique de Shafer pour définir un nouvel espace d'attributs invariants. En effet, si l'on considère un objet diélectrique inhomogène, toutes les coordonnées couleur C des pixels le constituant se situent sur une même ligne coïncidant avec le vecteur de *body* C_b associé au modèle de Shafer dans l'espace RGB (voir figure 8.1a). Les angles que forme ce vecteur C dans les plans bichromatiques $(C_i C_j)$ de l'espace RGB ($\arctan(C_i/C_j)$) sont des invariants photométriques, puisqu'ils ne varient pas en fonction de la position de C , et donc de l'intensité de l'illuminant. Basés sur ce principe, les attributs $c_1 c_2 c_3$ sont définis par :

$$c_1 = \arctan\left(\frac{R}{\max\{G, B\}}\right) \quad c_2 = \arctan\left(\frac{G}{\max\{R, B\}}\right) \quad c_3 = \arctan\left(\frac{B}{\max\{R, G\}}\right) \quad (8.13)$$

Ici aussi, ces valeurs sont bruitées dans le cas où chacune des composantes R, G et B est très faible.

Les invariants différentiels. Dans [Geusebroek *et al.*, 2001], les auteurs déterminent des invariants basés sur la dérivation de la radiance spectrale. Ainsi, en dérivant l'expression (8.8) par rapport aux longueurs d'onde λ (on obtient alors \mathcal{L}_λ), ils montrent que le rapport $F = \frac{\mathcal{L}_\lambda}{\mathcal{L}} = \frac{1}{\mathcal{R}_b(\lambda, p)} \frac{\partial \mathcal{R}_b(\lambda, p)}{\partial \lambda}$ est une caractéristique invariante par rapport à la direction d'observation, à l'orientation de surface, à la direction d'illumination, ainsi qu'à son intensité. En dérivant n fois la relation F par rapport aux coordonnées spatiales p , et m fois par rapport à la longueur d'onde λ , on obtient un ensemble d'invariants couleur qui peuvent être exprimés de la manière suivante :

$$F_{\lambda^m p^n} = \frac{\partial^n}{\partial p^n} \left\{ \frac{\mathcal{L}_{\lambda^m}}{\mathcal{L}} \right\} \text{ pour } m \geq 1, n \geq 0 \quad (8.14)$$

Un autre invariant différentiel est obtenu en considérant le cas d'un éclairage uniforme sur toute l'image et pour lesquels aucun effet d'ombre n'est observé. Dans ce cas, la luminance est constante en tout point p de l'image et $\mathcal{L} = \mathcal{E}\mathcal{R}_b(\lambda, p)$. Ainsi, tout rapport de la forme :

$$W_{\lambda^m p^n} = \frac{\mathcal{L}_{\lambda^m p^n}}{\mathcal{L}} \text{ pour } m \geq 0, n \geq 1 \quad (8.15)$$

exprime les changements de reflectance de l'objet indépendamment de l'intensité de l'illumination. Les quotients différentiels \mathcal{L}_{ij} sont calculés à partir d'un espace couleur gaussien [Geusebroek *et al.*, 2001]. Notons que les invariants différentiels sont bruités dans le cas de faibles reflectances \mathcal{L} .

Examinons à présent le cas d'un illuminant coloré.

8.2.2 Illuminant coloré

Contrairement au cas d'un illuminant blanc, la distribution spectrale d'une lumière colorée n'est pas constante suivant les longueurs d'onde. Néanmoins, il existe quelques invariants dédiés à ce cas de figure.

Les rapports de radiance. Si l'on considère que la reflectance diffuse du matériau \mathcal{R}_b est indépendante des longueurs d'onde de la lumière incidente $\mathcal{E}(\lambda)$, le modèle dichromatique devient :

$$C_i(p) = m_b(p)D_i(p) \int_{\lambda} S_i(\lambda)\mathcal{E}(\lambda, p)d\lambda \quad (8.16)$$

Posons $I_i(p) = \int_{\lambda} S_i(\lambda)\mathcal{E}(\lambda, p)d\lambda$. En supposant une illumination localement constante en deux pixels voisins p_1 et p_2 ($\mathbf{I}(p_1) = \mathbf{I}(p_2) = \mathbf{I}$), Gevers *et al.* [Gevers et Smeulders, 1997a] définissent un attribut couleur à la fois indépendant de la couleur de l'illuminant et insensible à la présence d'ombres dans l'image. Les coordonnées couleur de p_1 et p_2 écrites $C_i(p_k)$, pour $k = 1, 2$ sont exprimées par :

$$C_i(p_k) = m_b(p_k)D_i(p_k)I_i \quad (8.17)$$

Ainsi, le rapport de deux coordonnées couleur (C_i et C_j) au pixel p_k ne dépend plus du terme géométrique $m_b(p_k)$:

$$\frac{C_i(p_k)}{C_j(p_k)} = \frac{m_b(p_k)D_i(p_k)I_i}{m_b(p_k)D_j(p_k)I_j} = \frac{D_i(p_k)I_i}{D_j(p_k)I_j} \quad (8.18)$$

Les rapports décrits ci-après, qui mettent en jeu les couleurs aux deux pixels voisins p_1 et p_2 , ne dépendent plus de la couleur de l'illuminant \mathbf{I} :

$$\frac{C_i(p_1)C_j(p_2)}{C_j(p_1)C_i(p_2)} = \frac{D_i(p_1)I_i D_j(p_2)I_j}{D_j(p_1)I_j D_i(p_2)I_i} = \frac{D_i(p_1)D_j(p_2)}{D_j(p_1)D_i(p_2)} \quad (8.19)$$

Dans [Gevers et Smeulders, 1997a] est proposé l'espace d'attributs m_i , pour $i = 1..3$:

$$m_1 = \frac{R_1 G_2}{G_1 R_2}, \quad m_2 = \frac{B_1 R_2}{R_1 B_2}, \quad m_3 = \frac{G_1 B_2}{B_1 G_2} \quad (8.20)$$

Ces attributs sont donc indépendants de la couleur de l'illuminant mais également d'un changement d'angle d'observation, de l'orientation de surface et de l'intensité de l'illuminant. Dans le cas de deux pixels situés sur une région de couleur uniforme, les trois composantes m_i sont égales à 1. Notons cependant qu'il suffit qu'une seule coordonnée couleur (associée à l'un ou l'autre des pixels) soit proche de 0 pour que l'un des attributs m_1 , m_2 ou m_3 soit bruité.

Les attributs $\mathcal{C}_R, \mathcal{C}_B$. En se basant sur deux fortes approximations, à savoir que la caméra est équipée avec des capteurs à bande étroite² et que l'illuminant est *planckian**, de nouveaux invariants ont été proposés par [Gevers et Smeulders, 1997a, Finlayson *et al.*, 2000]. Dans ce cas de figure, les sensibilités spectrales des capteurs peuvent être exprimées à partir d'une fonction de Dirac, dont la sensibilité est réduite à une seule longueur d'onde $S_i(\lambda) = \delta(\lambda - \lambda_i)$ pour $i = \{R, G, B\}$. Dans ce cas, les coordonnées couleur sont exprimées par :

$$C_i(p) = \int_{\lambda} \delta(\lambda - \lambda_i) \mathcal{R}_b(\lambda, P) \mathcal{E}(\lambda, P) d\lambda \quad (8.21)$$

Du fait des propriétés de décalage de la fonction Dirac, on obtient $C_i(p) = I_i(p) D_i$. Ensuite, en modélisant l'illuminant $I_i(p)$ par une équation de Planck [Finlayson *et al.*, 2000], la réponse des capteurs devient :

$$C_i(p) = I(p) k_1 \lambda_i^{-5} e^{-\frac{k_2}{T \lambda_i}} D_i(p) \quad (8.22)$$

où k_1 et k_2 représentent des constantes physiques, T la température en Kelvin et enfin I l'amplitude d'illumination. En considérant le logarithme népérien $\ln C_i(p)$, la réponse du capteur peut être exprimée comme une somme de plusieurs termes. Une différence des logarithmes de deux coordonnées couleur $\ln C_i(p) - \ln C_j(p)$ permet alors d'annuler le terme dépendant de l'illumination I . Ce principe a été proposé par [Finlayson *et al.*, 2000], qui décrit les deux composantes \mathcal{C}_R et \mathcal{C}_G suivantes :

$$\mathcal{C}_R = \ln \frac{R}{G} = \ln(\lambda_R^{-5} D_R) - \ln(\lambda_G^{-5} D_G) + \frac{k_2}{T} \left(\frac{1}{\lambda_G} - \frac{1}{\lambda_R} \right) \quad (8.23)$$

$$\mathcal{C}_B = \ln \frac{B}{G} = \ln(\lambda_B^{-5} D_B) - \ln(\lambda_G^{-5} D_G) + \frac{k_2}{T} \left(\frac{1}{\lambda_G} - \frac{1}{\lambda_B} \right) \quad (8.24)$$

Ainsi, les valeurs $\mathcal{C}_R, \mathcal{C}_B$ ne dépendent plus des variations d'intensité de l'illuminant coloré.

Les invariants différentiels. En considérant dans la relation (8.8) que le spectre d'énergie ne varie pas dans l'image, $\mathcal{E}(P, \lambda)$ peut s'exprimer comme le produit d'un terme dépendant uniquement de la distribution spectrale d'énergie de l'illuminant $e(\lambda)$ et d'un terme $i(P)$ exprimant la variation d'intensité dans l'image due à la géométrie de la scène. Ainsi, dans le cas d'un objet mat et opaque, de coefficient de Fresnel proche de 0, la radiance s'exprime $\mathcal{L} = i(P) e(\lambda) \mathcal{R}_b(\lambda, P)$. L'attribut suivant [Geusebroek *et al.*, 2001] $N_{\lambda p} = \frac{\mathcal{L}_{\lambda p} \mathcal{L} - \mathcal{L}_{\lambda} \mathcal{L}_p}{\mathcal{L}^2}$ permet d'exprimer des changements de reflectance de l'objet indépendamment des changements

²Le domaine spectral de sensibilité des capteurs est très restreint.

d'angle d'observation, de l'orientation de la surface, de la direction d'éclairage, de son intensité ou de son spectre. Il en est de même pour l'ensemble des attributs couleur suivants :

$$N_{\lambda^m p^n} = \frac{\partial^{m+n-2}}{\partial \lambda^{m-1} \partial p^{n-1}} \left\{ \frac{\mathcal{L}_{\lambda p} \mathcal{L} - \mathcal{L}_{\lambda} \mathcal{L}_p}{\mathcal{L}^2} \right\} \quad (8.25)$$

Cette section a concerné les invariants couleur adaptés exclusivement aux matériaux diélectriques opaques, qu'ils soient exposés à un illuminant blanc ou coloré. Focalisons-nous à présent sur les invariants applicables aux matériaux conducteurs (spéculaires). Notons dès à présent que certains d'entre eux ne compensent pas totalement les changements d'éclairage.

8.3 Invariants adaptés aux matériaux conducteurs

À notre connaissance, la littérature ne propose pas d'invariants dédiés aux matériaux conducteurs, qui soient applicables dans le cas d'un illuminant coloré. Dans le cas d'un illuminant blanc, la couleur de la réflexion spéculaire est située sur l'axe des intensités du système RGB. Toutes les couleurs de l'objet observé (à la fois les couleurs issues de la réflexion diffuse et spéculaire) sont situées à l'intérieur du triangle formé par les deux composantes de réflexion C_s et C_b (voir la figure 8.1b). Le modèle dichromatique s'exprime alors par la relation (8.7). Les premiers attributs présentés ci-dessous se basent précisément sur cette hypothèse.

La teinte H. L'expression de la teinte, donnée par (2.14), dépend de différences de composantes couleur $C_i - C_j$. Or d'après (8.7), $C_i - C_j = I(p)m_b(p)(D_i - D_j)$. Cette différence permet d'annuler le terme spéculaire $I(p)m_s(p)Q$. Ensuite, en considérant une troisième composante couleur C_k , le rapport de deux différences de couleur annule l'influence de $I(p)m_b(p)$:

$$\frac{C_i - C_j}{C_i - C_k} = \frac{I(p)m_b(p)(D_i - D_j)}{I(p)m_b(p)(D_i - D_k)} = \frac{D_i - D_j}{D_i - D_k} \quad (8.26)$$

De telles valeurs sont donc invariantes vis-à-vis des spéularités, des variations d'éclairage et de la géométrie de la scène. De manière similaire, la teinte H est invariante à ces mêmes facteurs :

$$H = \arctan \left(\frac{\sqrt{3}(G - B)}{(R - G) + (R - B)} \right) = \arctan \left(\frac{\sqrt{3}(D_G - D_B)}{(D_R - D_G) + (D_R - D_B)} \right) \quad (8.27)$$

En effet, elle ne dépend que de la sensibilité des différents capteurs et de l'albedo de surface.

Les attributs $l_1 l_2 l_3$. À partir des propriétés du modèle dichromatique quant à la répartition des couleurs dans l'espace RGB, [Gevers et Smeulders, 1997a] proposent un nouvel ensemble d'attributs invariants à la géométrie de la scène. En effet, puisque toutes les couleurs d'un objet

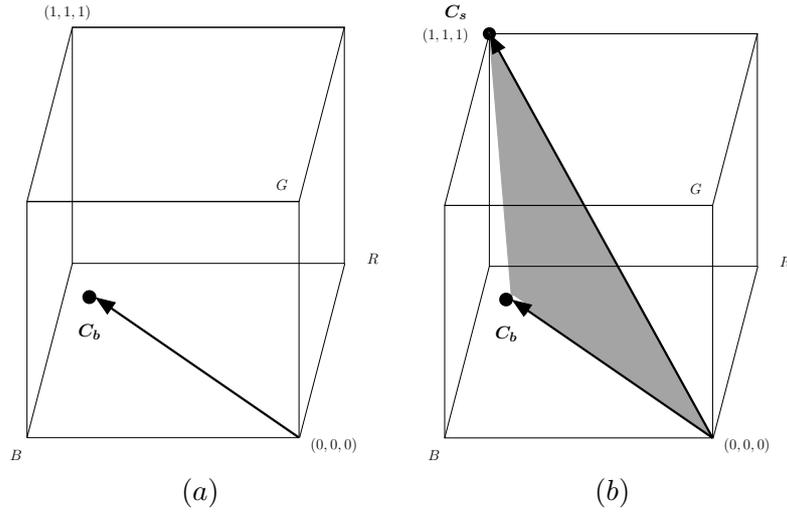


FIG. 8.1 – Modèle dichromatique (a) Représentation du modèle dichromatique dans le système RGB, dans le cas d'un matériau diélectrique inhomogène et opaque. (b) Représentation du modèle dichromatique dans le système RGB, dans le cas d'un illuminant blanc.

de couleur uniforme sont situées sur un plan formé par les couleurs de matériau C_b et d'illuminant C_s (cf. la figure 8.1b où ce plan est visualisé en gris), les directions de ce plan sur les trois composantes couleur forment un espace d'attributs $l_1 l_2 l_3$ photométriquement invariant :

$$l_1 = \frac{(R - G)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2} = \frac{(D_R - D_G)^2}{(D_R - D_G)^2 + (D_R - D_B)^2 + (D_G - D_B)^2} \quad (8.28)$$

$$l_2 = \frac{(R - B)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2} = \frac{(D_R - D_B)^2}{(D_R - D_G)^2 + (D_R - D_B)^2 + (D_G - D_B)^2} \quad (8.29)$$

$$l_3 = \frac{(G - B)^2}{(R - G)^2 + (R - B)^2 + (G - B)^2} = \frac{(D_G - D_B)^2}{(D_R - D_G)^2 + (D_R - D_B)^2 + (D_G - D_B)^2} \quad (8.30)$$

Les composantes de cet espace d'attributs ne dépendent donc plus que des valeurs D_i pour $i \in \{R, G, B\}$. Dans [Gevers *et al.*, 1998], les auteurs montrent l'invariance de la composante H et de l'espace $l_1 l_2 l_3$ dans une tâche de segmentation couleur. Or, les résultats obtenus sont médiocres dans les cas où la saturation est faible. Ainsi, il est préférable de ne pas considérer les valeurs de composantes associées à des valeurs de saturation trop faibles.

Les attributs $o_1 o_2$. Dans [Gevers et Stokman, 2003], les auteurs utilisent entre autre un espace bidimensionnel $o_1 o_2$ appelé espace couleur opposé :

$$o_1 = \frac{R - G}{2}, \quad o_2 = \frac{2B - R - G}{4} \quad (8.31)$$

Comme nous l'avons déjà précisé dans le cas de la teinte, une différence de deux composantes couleur permet d'annuler l'influence des spécularités en éliminant le terme constant $I(p)m_s(p)Q$ dans la relation (8.7), pourvu que l'illuminant soit blanc. Les attributs o_1 et o_2 sont donc indépendants du terme de reflectance spéculaire. Par contre, ils demeurent sensibles au terme géométrique $m_b(p)$ et aux variations d'intensité d'éclairage.

Les invariants différentiels. Dans [Geusebroek *et al.*, 2001], les auteurs prouvent l'invariance du rapport $\mathcal{H} = \mathcal{L}_\lambda / \mathcal{L}_{\lambda\lambda}$ par rapport à la position d'éclairage et d'observation, à l'intensité et au coefficient de Fresnel (\mathcal{L} étant donnée par (8.4) avec $I(\lambda, p) = I(\lambda)$, et \mathcal{L}_λ et $\mathcal{L}_{\lambda\lambda}$ sont ses dérivées première et seconde par rapport à λ). Ils montrent que cette valeur est équivalente à la teinte. Puisque les dérivées de \mathcal{H} par rapport à p et λ sont des invariants couleur, tous les attributs suivants sont également invariants :

$$\mathcal{H}_{\lambda^m p^n} = \frac{\partial^{m+n}}{\partial \lambda^m \partial p^n} \left\{ \arctan \left(\frac{\mathcal{L}_\lambda}{\mathcal{L}_{\lambda\lambda}} \right) \right\} \text{ pour } m, n \geq 0 \quad (8.32)$$

Les attributs invariants que nous présentons à présent ne se basent plus sur la relation (8.7).

Invariants photométriques. Comme nous l'avons vu dans le chapitre sur l'interaction lumière matière, la radiance provenant d'un point de la scène peut être décomposée en partie diffuse et partie spéculaire. De plus, chacune de ces composantes peut être exprimée comme le produit de deux termes, l'un dépendant des caractéristiques intrinsèques de l'objet (couleur, rugosité) et l'autre de la géométrie de la scène (direction d'éclairage, normale au point observé, direction d'observation). Ainsi, dans [Narasimhan *et al.*, 2003], les auteurs proposent d'exprimer la radiance de la manière suivante :

$$\mathcal{L}(\lambda, p) = \sum_{j=1}^n N_j(\lambda) G_j(p)$$

où les valeurs N dépendent exclusivement de la couleur du matériau et aucunement de la géométrie de la scène, et où les valeurs G ne dépendent que de la géométrie : forme de l'objet, éclairage, direction d'observation. Il s'agit finalement d'une extension du modèle dichromatique, dans le sens où il suppose que la réflectance peut se décomposer en plus de deux termes. Les auteurs montrent que, quel que soit le type de matériau, leur radiance s'exprime toujours ainsi, seul n étant modifié. En effet, dans le cas d'un modèle lambertien, $n = 1$, tandis que pour des matériaux rugueux et spéculaires, on aura $n = 2$. Dans ce cas, si \mathbf{E} est une matrice 3×3 comprenant les trois coordonnées couleur en trois points p_1 , p_2 et p_3 , elle peut s'écrire sous forme matricielle par :

$$\mathbf{E} = \mathbf{N}\mathbf{G}$$

où \mathbf{N} , de taille $3 \times n$, comprend les n coefficients N_j , et \mathbf{G} de taille $n \times 3$ les coefficients G_j pour $j = 1..n$. Les trois points p_1 , p_2 et p_3 peuvent correspondre au même point physique dans différentes images acquises sous différentes géométries d'éclairage. Ils peuvent également correspondre à trois points différents du même objet. Le rapport des déterminants de deux sous-matrices de \mathbf{E} de taille 2×2 est indépendant vis-à-vis de la géométrie de la scène. Soit E_j^i , la composante couleur i du point p_j . Par exemple, le rapport suivant :

$$\frac{E_1^R E_2^G - E_2^R E_1^G}{E_1^R E_3^G - E_3^R E_1^G} = \frac{N_1^R N_2^G - N_2^R N_1^G}{N_1^R N_3^G - N_3^R N_1^G}$$

constitue un attribut invariant dans le sens où il ne dépend plus que des termes N_j relatifs à la couleur du matériau. Ceci est vrai pour n'importe quelle sous-matrice de \mathbf{E} . Néanmoins,

mettons certaines réserves quant à l'utilisation de cet invariant. En effet, il s'avère très bruité dès lors que le changement d'éclairage n'est pas suffisant pour que l'aspect de l'objet soit modifié, et ainsi $E_3^R E_1^G = E_1^R E_3^G$.

8.4 Conclusion

Contrairement au cas des images de luminance où des approximations locales des variations d'illumination sont à ce jour les seules possibilités de compenser les variations d'illumination, les images couleur apportent des solutions intéressantes, par la définition d'invariants couleur valables globalement, en chaque pixel de l'image. Dans ce chapitre nous avons fait le bilan des attributs couleurs invariants aux changements d'intensité d'illumination et à la géométrie de la scène (positions relatives objet-éclairage-capteur). Nous n'avons pas considéré le problème de la constance couleur, qui constitue à notre sens un domaine d'étude en soit.

Toutefois, des travaux semblent encore nécessaires pour l'obtention d'attributs qui soient invariants à tout type de modification des conditions d'acquisition. Notamment, il n'existe pas d'invariants adaptés aux matériaux diélectriques dans le cas où l'illuminant est coloré. De plus, nous avons souligné l'aspect bruité de la plupart de ces attributs, généralement lorsque les luminances sont faibles. Il s'agit à notre sens de leur principale limitation. Dans la suite du document, nous nous attachons à utiliser les propriétés d'invariance couleur pour améliorer le suivi de points d'intérêt dans les images couleur.

SUIVI DE POINTS COULEUR ROBUSTE AUX VARIATIONS D'ILLUMINATION

Ce chapitre s'intéresse au suivi de points dans les séquences d'images en couleur. Nous améliorons la robustesse du suivi vis-à-vis des variations d'illumination, d'une part en étendant les modèles photométriques locaux exposés au chapitre 6 au cas des images couleur, d'autre part en exploitant les invariants couleur décrits dans le chapitre 8. En outre, puisque certains invariants très pertinents ne compensent pas la totalité des variations d'illumination susceptibles de se produire entre deux images d'une même séquence, nous proposons d'améliorer leur utilisation par un modèle d'illumination local.



L'intérêt des attributs de couleur en suivi d'objets n'est plus à démontrer à l'heure actuelle. Nombreuses sont les applications en suivi de personnes ou de groupes de personnes en mouvement. Cela s'explique notamment par l'évolution très significative ces dernières années de l'automatisation des systèmes de surveillance. Il existe même des applications de surveillance des interactions entre personnes [MacKenna *et al.*, 2000], afin de détecter des situations inhabituelles dans les scènes observées. Soulignons également l'existence d'applications dans le domaine sportif, par exemple le suivi de joueurs [Vandenbroucke, 2000] et de ballon sur les stades de football. Ce genre d'application illustre bien l'intérêt des attributs couleur, qui constituent un moyen très efficace de distinguer les équipes entre elles. De même, les applications d'interaction homme-machine basées sur la vision nécessitent le suivi de mouvements humains [Wu et Huang, 2000] ou la détermination des expressions du visage [Stern et Efros, 2005, Spors et Rabenstein, 2001]. Il s'agit alors de capter les expressions humaines et de les analyser. Le lecteur intéressé par ce type d'applications, pourra se référer à la revue exhaustive en analyse de dynamique humaine [Wang et Singh, 2003].

Dans tous ces domaines d'applications, l'utilisation de la couleur est très pertinente puisqu'elle apporte énormément d'information discriminante utile au suivi.

Les attributs d'invariance couleur exposés dans le chapitre 8 font l'objet d'un grand intérêt dans le cadre du suivi d'objet ou du flot optique mais n'ont, à notre connaissance, jamais été exploités en suivi de primitives point.

Les invariants les plus couramment utilisés en vision par ordinateur sont sans conteste les attributs RGB normalisés par la norme L_1 , qui ont été définis par les relations (2.10). Citons par exemple les travaux décrits dans [Oliver *et al.*, 1997, Soriano *et al.*, 2000, MacKenna *et al.*, 2000]. L'espace couleur de Teinte-Luminance-Saturation, HSV, est également largement utilisé pour le suivi [Tadjine et Joubert, 2003, Sigal *et al.*, 2000, Jones et Rehg, 1999]. Dans [Tadjine et Joubert, 2003], l'originalité de l'approche réside dans l'utilisation d'attributs couleur invariants, tels que les espaces $l_1l_2l_3$ et $c_1c_2c_3$, dans le calcul des gradients couleur. Ces espaces permettent d'obtenir des valeurs de gradients faibles au niveau des contours associés aux ombres. Les invariants couleur $c_1c_2c_3$ ont également été utilisés dans un contexte de suivi par filtrage de Kalman [Nguyen et Smeulders, 2002].

Contrairement au domaine du suivi d'objets, pour lequel l'utilisation de la couleur s'avère largement répandue, le suivi de points caractéristiques dans les images couleur fait l'objet de très peu de publications. Citons toutefois les travaux de Heigl, qui a étendu le suivi de points classique KLT [Lucas et Kanade, 1981, Tomasi et Kanade, 1991] au cas des images en couleur [Heigl *et al.*, 1999]. À notre sens, le peu d'intérêt suscité par le suivi de points couleur est du aux temps de calcul qu'implique la prise en compte d'images vectorielles. Pourtant, la pertinence de l'information couleur a notamment été démontrée dans le cadre de la mise en correspondance d'images [Gouet *et al.*, 2000] qui se base sur les invariants différentiels de Schmid [Schmid et Mohr, 1997] et rend la mise en correspondance invariante aux changements d'illumination. De plus, l'utilisation de la couleur doit permettre d'extraire (et donc suivre) un nombre plus important de points par rapport aux images de luminance, dans le sens où l'on dispose d'une information supplémentaire de chrominance.

L'utilisation des attributs d'invariance couleur doit permettre d'améliorer la robustesse du suivi vis-à-vis des changements d'illumination. Nous proposons ainsi trois approches permettant d'améliorer la robustesse du suivi dans les images couleur.

- Dans un premier temps, une modélisation *locale* des variations d'illumination, telle que celles proposées dans le chapitre 6, est intégrée dans la tâche de suivi.
- Ensuite, quelques uns des invariants couleur décrits dans le chapitre 8, qui permettent une correction *globale* des changements d'illumination, sont exploités.
- Pour finir, puisque la plupart des invariants couleur ne compense pas correctement la globalité des variations d'illumination, en particulier celles dues à la présence de spécularités, nous proposons d'améliorer leur utilisation grâce à l'ajout d'une correction *locale*.

Ce chapitre s'articule donc ainsi. La section 9.1 décrit brièvement la technique de suivi classique, dite KLT, étendue au cas des images en couleur. La première approche proposée, basée sur l'utilisation d'un modèle photométrique, est explicitée dans la section 9.2. La section 9.3 exploite quelques invariants couleur dans le cadre du suivi de points. La méthode de suivi utilisant conjointement les invariants couleur et la modélisation photométrique est abordée dans la section 9.4. Enfin, une étude comparative des différentes techniques de suivi est établie dans la

section 9.5. Différentes expérimentations sont considérées, sur des scènes présentant des objets spéculaires ou lambertiens, soumis ou non à des variations d'illumination.

9.1 Suivi de points couleur

La principale difficulté du suivi dans des images couleur réside dans la prise en compte du caractère vectoriel qui lui est inhérent. Une solution consisterait à adopter une méthode marginale, en effectuant le calcul de la déformation géométrique pour chacune des coordonnées couleur de manière indépendante, pour ensuite moyenner par exemple le modèle afin d'en avoir une approximation. À notre sens, chacun des plans couleur apporte seul moins d'information pertinente que l'image de luminance. Par exemple, un des plans couleur peut être nul en chaque point, et engendrer un mauvais calcul de mouvement. Ainsi, il est probable que le suivi sur les trois plans indépendamment mène finalement à un résultat moins précis qu'avec une image de luminance.

Ainsi, le caractère vectoriel de la couleur doit être pris en compte dans les techniques de suivi. Soit $\mathbf{f}(p)$ le vecteur image, à K composantes, au point p à l'instant k et $\mathbf{g}(p)$ le vecteur image au point p' à l'instant k' . Dans ce cas, le modèle de mouvement $\delta(p, \mathbf{A})$, qui a été décrit dans la section 7.2.2.2, peut se calculer par minimisation de la fonction de coût suivante :

$$\epsilon_1(\mathbf{A}) = \frac{1}{2} \sum_{m \in \mathcal{W}} \|\mathbf{f}(m) - \mathbf{g}(\delta(m, \mathbf{A}))\|^2 \quad (9.1)$$

où \mathcal{W} est une fenêtre d'intérêt de taille $\mathcal{N} \times \mathcal{N}$ et m est un point situé dans cette fenêtre \mathcal{W} . L'obtention de \mathbf{A} s'effectue sur le même principe que dans les images de luminance (voir le paragraphe 7.2.2.1). On suppose que $\mathbf{A} = \hat{\mathbf{A}} + \mathbf{a}$ où \mathbf{a} exprime une petite variation autour d'une estimation $\hat{\mathbf{A}}$ de \mathbf{A} puis on développe au 1^{er} ordre $\mathbf{g}(\delta(m, \mathbf{A}))$ en série de Taylor autour de $\hat{\mathbf{A}}$. En négligeant les termes du second ordre, nous obtenons finalement :

$$\mathbf{g}(\delta(m, \mathbf{A})) = \mathbf{g}(\delta(m, \hat{\mathbf{A}})) + \mathbf{G}_g(\delta(m, \hat{\mathbf{A}})) J_\delta^{\mathbf{A}} \mathbf{a} \quad (9.2)$$

où \mathbf{G}_g est la matrice jacobienne de l'image g par rapport à x et y . En injectant (9.2) dans (9.1), on aboutit finalement à un système linéaire en \mathbf{a} :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V}_C \mathbf{V}_C^T \right) \mathbf{a} = \frac{1}{K} \sum_{m \in \mathcal{W}} \left(\sum_i^K (f_i(m) - g_i(\delta(m, \hat{\mathbf{A}}))) \mathbf{V}_i \right) \quad (9.3)$$

où le vecteur \mathbf{V}_i est exprimé dans le cas d'un modèle affine par :

$$\mathbf{V}_i = [g_x^i, g_y^i, xg_x^i, xg_y^i, yg_x^i, yg_y^i]^T \quad \forall i = 1 \dots K \quad (9.4)$$

avec g_x^i et g_y^i les dérivées premières en m du plan image i . Quant au vecteur \mathbf{V}_C , il s'écrit :

$$\mathbf{V}_C = [g_x, g_y, xg_x, xg_y, yg_x, yg_y]^T \quad (9.5)$$

avec :

$$g_x = \frac{1}{K} \sum_i^K g_x^i \quad g_y = \frac{1}{K} \sum_i^K g_y^i \quad (9.6)$$

Le suivi de points d'intérêt dans les images couleur a été défini dans [Heigl *et al.*, 1999]. Comme pour l'approche classique du suivi de points dans les images de luminance, présentée dans la section 7.2.2.1, cette technique n'est robuste à aucun changement d'illumination. Puisqu'elle se base sur le modèle de conservation des luminances (voir la section 6.2.1) elle suppose implicitement que les surfaces de la scène sont lambertiennes, que l'éclairage ne varie pas, que ce soit en intensité en couleur ou en position, et que le gain de la caméra reste constant entre les instants k et k' .

La section suivante expose une première contribution de ce chapitre, qui correspond à l'extension de l'approche développée dans le chapitre 7 pour les images de luminance au cas des images en couleur.

9.2 Suivi de points couleur basé sur une approximation photométrique locale

Dans le chapitre 7, nous avons montré la pertinence de l'approche de suivi proposée dans le cas d'images de luminance. D'après les expérimentations mises en œuvre, elle s'avère plus performante (dans le cadre du suivi de points d'intérêt) que les approches qui supposent un modèle d'illumination affine [Jin *et al.*, 2001, Tommasini *et al.*, 1999]. Dans le cas d'images couleur, les variations d'illumination peuvent être exprimées comme dans la relation (6.32). En exploitant ce modèle dans le cadre du suivi de points couleur, le calcul des paramètres du modèle de mouvement revient à minimiser le critère suivant :

$$\epsilon_2(\mathbf{A}) = \frac{1}{2} \sum_{m \in \mathcal{W}} \| \mathbf{f}(m) - \mathbf{g}(\delta(m, \mathbf{A})) + \mathbf{U}^T \mathbf{B} \|^2 \quad (9.7)$$

où $\mathbf{U} = (x - x_p, y - y_p, 1)^T$ et \mathbf{B} est la matrice comprenant les termes du modèle photométrique $\mathbf{B} = (\alpha, \beta, \gamma)$. En admettant un petit déplacement \mathbf{a} et en approximant $\mathbf{g}(\delta(m, \mathbf{A}))$ par un développement en série de Taylor, comme dans (9.2), le calcul de \mathbf{a} s'obtient par minimisation du système linéaire suivant :

$$\left(\sum_{m \in \mathcal{W}} \mathbf{V}_P \mathbf{V}_P^T \right) \mathbf{a} = \frac{1}{K} \sum_{m \in \mathcal{W}} \left(\sum_i^K \left(f_i(m) - g_i(\delta(m, \hat{\mathbf{A}}) - \mathbf{U}^T \mathbf{B}_i) \right) \mathbf{V}_P^i \right) \quad (9.8)$$

avec $\mathbf{U} = (x - x_p, y - y_p, 1)^T$, $\mathbf{B}_i = (\alpha_i, \beta_i, \gamma_i)^T$. Dans le cas d'un modèle affine de mouvement, les vecteurs \mathbf{V}_P^i s'écrivent :

$$\mathbf{V}_P^i = [\mathbf{V}_i^T, \mathbf{U}^T]^T \quad \forall i = 1 \dots K \quad (9.9)$$

Le vecteur \mathbf{V}_P s'exprime quant à lui :

$$\mathbf{V}_P = [\mathbf{V}_C \mathbf{V}_C^T, \mathbf{U}^T]^T \quad (9.10)$$

où le vecteur V_C est décrit par (9.5).

Dans le cas où l'illuminant est blanc, nous avons vu dans le paragraphe 6.3.1, que chacun des paramètres photométriques exprimés dans un espace couleur RGB (α , β et γ) était scalaire.

Après avoir proposé une méthode de suivi de points basée sur une approximation locale des variations d'illumination, nous présentons dans la prochaine section des méthodes de correction globale, par utilisation des invariants couleur.

9.3 Suivi couleur par utilisation d'invariants

Le chapitre 8 a recensé un certain nombre d'attributs couleur permettant d'assurer une invariance vis-à-vis des changements d'illumination. Nous justifions ici le choix de quelques uns d'entre eux dans le cadre du suivi de points d'intérêt. La technique de suivi décrite dans la section 9.1, qui s'applique aux images vectorielles, est également adaptée à ce genre d'attributs.

9.3.1 Attributs existants

Parmi les attributs invariants présentés dans le chapitre 8, nous ne considérons pas les attributs différentiels dans le sens où les dérivées mises en jeu peuvent s'avérer bruitées. Ensuite, les attributs C_R et C_B reposent sur l'hypothèse selon laquelle l'illuminant est planckian, hypothèse limitative dans le sens où tous les types d'illuminants ne sont pas pris en compte. Nous avons vu que les attributs o_1o_2 (voir la section 8.3) sont invariants vis-à-vis des spécularités dans le cas particulier d'un illuminant blanc. De plus, ce ne sont pas les invariants les plus adaptés dans le sens où ils ne permettent pas de compenser les variations d'intensité d'éclairage. Pour finir, nous ne sélectionnons pas non plus les rapports de radiance présentés dans la section 8.2.2, qui supposent l'hypothèse d'une illumination localement constante.

Finalement, nous avons choisi de se focaliser sur les invariants normalisés (par les normes L_1 et L_2), les attributs $c_1c_2c_3$, la teinte et $l_1l_2l_3$. Certes, ces attributs ne sont applicables qu'aux objets diélectriques homogènes et opaques, mais ont le précieux avantage de ne dépendre ni de la géométrie d'éclairage et d'observation, ni de l'intensité de l'illumination. Quant aux composantes $l_1l_2l_3$ et H (voir la section 8.3), elles sont non seulement invariantes vis-à-vis de la géométrie de la scène et de l'intensité d'éclairage mais également vis-à-vis de l'occurrence de spécularités. Dans ce sens, ces attributs sont tout à fait intéressants. Toutefois, les composantes $l_1l_2l_3$ sont des fonctions bijectives, c'est-à-dire que deux combinaisons de R , G et B différentes peuvent aboutir à deux valeurs identiques des attributs $l_1l_2l_3$. Nous proposons de les modifier légèrement pour aboutir aux attributs $A_1A_2A_3$ décrits dans le paragraphe qui suit.

9.3.2 Attributs $A_1A_2A_3$

D'après l'étude du paragraphe 8.3, tout rapport de différences de composantes couleur est invariant vis-à-vis des variations d'illumination et des spéularités. En se basant sur ce constat, nous proposons ici les attributs $A_1A_2A_3$, exprimés comme suit :

$$\begin{aligned} A_1 &= \frac{R - G}{|R - G| + |G - B| + |B - R|} \\ A_2 &= \frac{R - B}{|R - G| + |G - B| + |B - R|} \\ A_3 &= \frac{G - B}{|R - G| + |G - B| + |B - R|} \end{aligned} \quad (9.11)$$

(9.12)

Ces attributs sont un bon compromis par rapport à la teinte et $l_1l_2l_3$. D'une part, le défaut de surjectivité est réduit, du fait que l'on n'introduit pas de valeur absolue au numérateur. D'autre part il permet de conserver une information tridimensionnelle, contrairement à la teinte.

Malheureusement, ces attributs invariants (H , $l_1l_2l_3$ et $A_1A_2A_3$) sont bruités lorsque leur dénominateur est faible. Cela se produit pour des luminances faibles, mais également pour de faibles saturations des couleurs ($R \simeq G \simeq B$). Soulignons que les invariants normalisés (par les normes L_1 et L_2) et les attributs $c_1c_2c_3$ ne sont quant à eux bruités que dans le cas de luminances faibles.

L'annexe D compare les invariants choisis pour différentes variations d'illumination et montre combien ces attributs permettent de réduire significativement leur influence en comparaison aux coordonnées couleur classiques. Cette annexe montre également que les attributs H , $l_1l_2l_3$ et $A_1A_2A_3$ peuvent réduire la séparabilité entre certaines couleurs. C'est d'autant plus vrai dans le cas de la teinte, qui est définie par un seul plan couleur.

Finalement, nous écartons la teinte et les attributs $l_1l_2l_3$. Nous proposons d'exploiter les attributs invariants dédiés aux matériaux lambertiens (qui n'annulent pas l'influence des spéularités) en les améliorant par l'ajout d'un modèle photométrique local dans la fenêtre d'intérêt centrée autour du point à suivre. Nous avons en effet montré dans le chapitre 7, que le modèle décrit par (6.3.1) s'avère tout à fait approprié pour compenser les variations de spéularité.

9.4 Suivi couleur par utilisation d'invariants et modélisation locale des variations de spéularité

Comme nous l'avons vu dans la section précédente, certains invariants couleur ne prennent pas en compte l'apparition de spéularités (espaces normalisés L_1 , L_2 , espace $c_1c_2c_3$), mais proposent une bonne invariance vis-à-vis des changements d'éclairage (voir les résultats de comparaison dans l'annexe D). Il est possible d'exprimer la part de variation d'illumination qui n'est pas prise en compte par ces invariants. Nous avons vu au chapitre 8 (voir la relation (8.7) page 183), que dans le cas d'un illuminant blanc (ou lorsqu'une balance des blancs est

effectuée), le modèle dichromatique exprime la valeur d'un vecteur RGB au point p de la manière suivante :

$$\mathbf{C}(p) = I(p)m_b(p)\mathbf{D}(p) + I(p)m_s(p)Q \quad (9.13)$$

où $I(p)$ correspond à l'intensité de l'illuminant, $\mathbf{D}(p)$ dépend à la fois de la composante de réflexion diffuse du matériau et de la sensibilité des capteurs, et où Q dépend de la sensibilité du capteur. Les termes m_s et m_b sont associés quant à eux à la géométrie de la scène. Posons $K_1(p) = I(p)m_b(p)$ et $K_2 = I(p)m_s(p)Q$. K_1 et K_2 sont identiques sur les trois plans couleur. En prenant comme exemple les attributs de norme L_1 , exprimons la part de variation d'illumination qui n'est pas prise en compte par ces attributs.

Notons r_{spec} la composante R normalisée par la norme L_1 dans le cas d'un matériau spéculaire. Elle est donnée par :

$$r_{spec}(p) = \frac{K_1(p)D_R(p) + K_2(p)}{K_1(p)(D_R(p) + D_G(p) + D_B(p)) + 3K_2(p)} \quad (9.14)$$

Dans le cas d'un matériau lambertien, $K_2 = 0$. En notant r_{dif} la composante normalisée associée à ce genre de matériau, l'influence des spécularités qui n'est pas compensée par la normalisation est donnée par $d_{spec}(p) = r_{spec}(p) - r_{dif}(p)$ de sorte que :

$$d_{spec}(p) = \frac{Num}{Den} = \frac{K_2(p)(D_G(p) + D_B(p) - 2D_R(p))}{K_1(p)(D_R(p) + D_G(p) + D_B(p))^2 + 3K_2(p)(D_R(p) + D_G(p) + D_B(p))} \quad (9.15)$$

La différence $\nu = Num - Den$ nous permet de comparer $d_{spec}(p)$ par rapport à 1. On a :

$$\nu = - [2K_2(p)(D_B(p) + D_G(p)) + 5K_2(p)D_R(p) + K_1(p)(D_R(p) + D_G(p) + D_B(p))^2] \quad (9.16)$$

Soulignons que les valeurs K_1 et K_2 dépendent de l'intensité. On peut dire sans perte de généralité que $K_1 \gg 1$ et $K_2 \gg 1$. D'autre part, D_R , D_G et D_B sont des valeurs positives. On peut en déduire que $\nu \ll 0$, et par conséquent que $d_{spec}(p) \ll 1$. D'autre part, il est légitime de penser que la part de changement d'illumination qui n'est pas prise en compte par l'invariant est faible par rapport à la valeur de l'invariant.

Nous supposons que $d_{spec}(p)$ peut être décrite par une fonction de classe C^1 de sorte qu'elle puisse être développée en série de Taylor d'ordre 1 dans un voisinage \mathcal{W} autour de p . Soit un point m appartenant à \mathcal{W} . L'image \mathbf{f} d'invariant à l'instant k et l'image d'invariant \mathbf{g} à l'instant k' sont alors reliées par :

$$\mathbf{g}(m') = \mathbf{f}(m) + \boldsymbol{\alpha}(x - x_p) + \boldsymbol{\beta}(y - y_p) + \boldsymbol{\gamma} \quad (9.17)$$

Cette analyse peut également être faite à partir des autres attributs invariants valables pour les matériaux lambertiens.

Ainsi, l'approche décrite dans la section 9.2 peut être exploitée en utilisant des invariants couleur (norme L_1 , norme L_2 , $c_1c_2c_3$) où il s'agit de calculer les paramètres du modèle de mouvement par minimisation de la fonction de coût ϵ_2 (équation (9.7)).

La prochaine section permet, par le biais de résultats expérimentaux, de comparer les différentes approches considérées dans ce chapitre : ces méthodes utilisent les composantes RGB ou les invariants couleur, avec ou sans modélisation photométrique locale.

9.5 Résultats expérimentaux

Dans un premier temps, nous explicitons la méthode de détection de points employée. Les différentes méthodes de suivi de points utilisant les invariants avec calcul du modèle photométrique (minimisation de ϵ_2) et sans calcul du modèle (ϵ_1) sont ensuite comparées.

9.5.1 Détection de points

La détection des points par la méthode de Harris [Harris et Stephens, 1988] dans les images d'intensité a été étendue au cas de la couleur. Elle nécessite alors le calcul de la matrice suivante [Gouet *et al.*, 2000] :

$$\begin{pmatrix} \sum_{\mathcal{W}}(R_x^2 + G_x^2 + B_x^2) & \sum_{\mathcal{W}}(R_x R_y + G_x G_y + B_x B_y) \\ \sum_{\mathcal{W}}(R_x R_y + G_x G_y + B_x B_y) & \sum_{\mathcal{W}}(R_y^2 + G_y^2 + B_y^2) \end{pmatrix} \quad (9.18)$$

$R_x, B_x, G_x, R_y, B_y, G_y$ étant les dérivées spatiales en x et y de l'image RGB . Après calcul des valeurs propres de cette matrice, λ_1 et λ_2 , le point est accepté si : $\min(\lambda_1, \lambda_2) > S_\lambda$, S_λ étant un seuil dont la valeur influe sur le nombre de points d'intérêt sélectionnés dans l'image initiale. Par ailleurs, cette technique pourrait être adaptée à d'autres espaces couleur, mais également aux invariants couleur. Néanmoins, en fonction des attributs utilisés, les points sélectionnés seraient très probablement différents. Or, notre étude nécessite de détecter les mêmes points pour toutes les approches dans le but de les comparer. En outre, afin de prendre en compte la particularité des attributs invariants, pour la plupart non définis pour les couleur non saturées, nous améliorons la détection en rajoutant un poids devant chaque composante de la matrice (9.18), traduisant la saturation S des pixels exprimée par (2.15) (voir la section 2.3.3 page 38).

9.5.2 Comparaison des méthodes de suivi

Nous comparons le suivi de points par utilisation des invariants choisis dans la section 9.3 à savoir avec la méthode de suivi classique dans RGB . Dans les deux cas, les approches décrites dans les sections 9.1 et 9.2 sont considérées. Nous comparons également ces approches aux techniques dédiées aux images de luminance [Gouiffès *et al.*, 2005], décrite respectivement dans les sections 7.2.2.1 et 7.4. Afin de simplifier les commentaires liés aux expérimentations, nous notons C l'approche de suivi classique dans les images de luminance, RGB l'approche classique pour les images couleur, et $L_1, L_2, c_1 c_2 c_3$ ainsi que $A_1 A_2 A_3$ les approches exploitant les invariants couleur. Enfin, les approches basées sur les composantes $Comp$ ($C, RGB, L_1, L_2, c_1 c_2 c_3$ ou $A_1 A_2 A_3$), sont notées $Comp + P_3$.

Ces différentes approches sont comparées sur des séquences d'images montrant à la fois des déformations géométriques et photométriques. Pour chacune des méthodes, nous sélectionnons les mêmes points, comme expliqué dans la section 9.5.1 et le suivi est basé sur le calcul du modèle de mouvement affine décrit dans la section 7.2.2.2 page 152. Un point est rejeté dès que ses résidus dépassent un seuil $S_{conv} = \mathcal{N}^2 E_{moy}^2$, où \mathcal{N} est la taille de \mathcal{W} , en pixels, et E_{moy} la variation d'intensité tolérée entre f (image initiale) and g (image courante), en pourcentage. Dans ces expérimentations, $E_{moy} = 15\%$. Ici, étant donné le caractère bruité des images (elles

sont acquises avec une caméra à trames entrelacées), nous considérons des tailles de fenêtre $\mathcal{N} = 11$ et $\mathcal{N} = 15$. Dans chaque cas, la séquence est jouée de la première image à la dernière puis de la dernière à la première, afin de pouvoir observer la symétrie des courbes d'évolution des résidus. Deux critères d'évaluation sont considérés :

- la *robustesse* du suivi, par le biais du nombre de points suivis par chaque approche ;
- la *précision* du suivi, c'est-à-dire les résidus moyens obtenus.

Nous étudions l'efficacité du suivi vis-à-vis des variations de spécularité puis des variations d'éclairage.

9.5.2.1 Apparition de spécularités

Séquence 1

Des images de cette séquence sont montrées sur la figure 9.1a. Elle représente un tableau sous verre, que l'on déplace par rapport à la caméra. Les changements d'orientation de l'objet par rapport aux sources d'éclairage (lumière du jour et tubes fluorescents), des réflexions spéculaires apparaissent. Aucune variation d'éclairage n'est provoquée.

Le nombre de points suivis par chaque technique (sur 10 points sélectionnés) est recensé dans le tableau 9.1b, suivant l'approche utilisée (9.1) (minimisation de ϵ_1) ou (9.7) (minimisation de ϵ_2). En considérant la première technique (ϵ_1), l'utilisation des invariants normalisés L_2 et $c_1 c_2 c_3$ permet d'assurer une meilleure robustesse du suivi par rapport aux autres approches. L'utilisation du modèle photométrique (méthode ϵ_2), permet d'améliorer la robustesse, par rapport à l'approche ϵ_1 . Plus particulièrement, l'association des composantes RGB et du modèle photométrique s'avère robuste quel que soit \mathcal{N} . Elle permet notamment de suivre plus de points que la méthode développée pour les images de luminance ($C + P_3$). Enfin, pour $\mathcal{N} = 15$, la méthode $L_2 + P_3$ permet d'obtenir d'aussi bons résultats que $RGB + P_3$. Notons également que l'association d'un modèle photométrique et d'invariants couleur est plus pertinente que l'utilisation d'attributs invariants pour matériaux spéculaires tels que $A_1 A_2 A_3$.

La figure 9.1c montre les résidus de convergence obtenus avec la méthode classique RGB et les deux approches les plus pertinentes ($RGB + P_3$) et la méthode conjointe normalisation L_2 et modèle local ($L_2 + P_3$). Dès la 50^e itération environ, la méthode classique ne suit plus aucun point tandis que les techniques avec modélisation locale ($RGB + P_3$ et $L_2 + P_3$) obtiennent de faibles résidus. Finalement, l'approche $L_2 + P_3$ compense plus significativement les variations spéculaires par l'utilisation conjointe des invariants et d'une modélisation de l'illumination.

Séquence 2

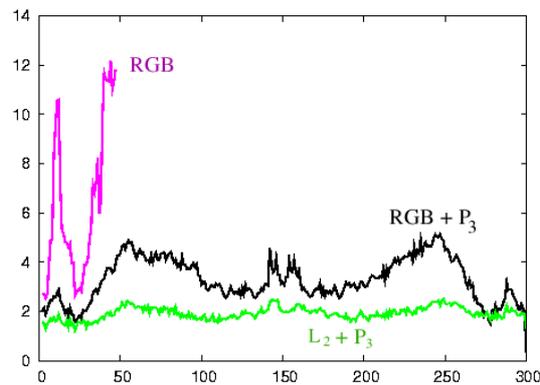
La deuxième séquence (figure 9.2a) représente une boîte cartonnée dont la surface est lisse et spéculaire. Les couleurs sont moins saturées que pour la séquence précédente. La caméra et l'éclairage sont fixes et l'objet est déplacé. Là encore aucune variation d'éclairage n'est provoquée, mais des spécularités sont engendrées par le mouvement de l'objet vis-à-vis de l'éclairage et de la caméra. Initialement, 36 points sont détectés. Comme le prouvent les résultats quant aux nombres de points suivis (tableau 9.2b), l'approximation d'un modèle photométrique améliore significativement la robustesse du suivi. De plus, les résultats s'avèrent meilleurs avec les composantes RGB qu'avec les invariants, quelle que soit \mathcal{N} . Pour $\mathcal{N} = 15$, la robustesse obtenue



(a)

Attributs	Méthode ϵ_1 (éq. (9.1)) <i>Comp</i>		Méthode ϵ_2 (éq. (9.7)) <i>Comp + P₃</i>	
	$\mathcal{N}=11$	$\mathcal{N}=15$	$\mathcal{N}=11$	$\mathcal{N}=15$
NdG	0	0	2	6
RGB	0	0	5	7
norme L_1	0	0	1	3
norme L_2	2	3	2	7
$c_1c_2c_3$	1	3	1	4
$A_1A_2A_3$	0	0	-	-

(b)



(c)

FIG. 9.1 – Séquence 1. (a) Deux images de la séquence. (b) Nombre de points suivis (sur 10). (c) Résidus de convergence des méthodes RGB , $RGB + P_3$ et $L_2 + P_3$, pour $\mathcal{N} = 15$.

avec les invariants $c_1c_2c_3$ et L_2 est également satisfaisante, puisqu'ils suivent un grand nombre de points. La figure 9.2c montre les résidus obtenus avec $RGB + P_3$, $c_1c_2c_3 + P_3$ et enfin RGB . Cette dernière technique perd tous les points considérés dès la 25^e itération. Les résidus les plus faibles sont obtenus avec l'approche $L_2 + P_3$, bien que le nombre de points suivis soit moins important pour cette méthode (voir tableau 9.2c). Malgré la technique de sélection des points mise en œuvre, l'utilisation des invariants se heurte là encore au problème de l'apparition de bruit pour des couleurs peu saturées. Les variations d'illumination influent sur cette satura-

tion au cours de la séquence. Cela justifierait d'utiliser une méthode de détection plus adaptée, c'est-à-dire dédiée aux attributs couleur utilisés.

Ainsi, la robustesse du suivi de points vis-à-vis de l'apparition de spécularités est améliorée par l'utilisation de la couleur. L'utilisation des invariants couleur apporte une bonne robustesse mais l'ajout d'une modélisation photométrique s'avère encore plus pertinente.

Intéressons-nous à présent au cas des variations d'intensité d'éclairage.

9.5.2.2 Variations d'intensité d'éclairage

Séquence 3

La troisième séquence (figure 9.3a) montre différents objets de divers matériaux, éclairés par des tubes fluorescents et un éclairage halogène. Seule la caméra est déplacée. La scène met en jeu des surfaces spéculaires mais également des variations d'illumination causées par l'apparition d'ombres projetées sur les objets (liées au passage d'un opérateur entre l'une des sources d'éclairage et la scène). Le tableau 9.3b recense le nombre de points suivis sur 39 sélectionnés initialement. Le suivi avec correction globale (par utilisation d'invariants) et le suivi avec correction locale ($RGB + P_3$) améliorent très sensiblement la robustesse du suivi. Pour $\mathcal{N}=11$, l'utilisation conjointe des invariants et de la correction locale n'améliore pas les résultats, contrairement au cas $\mathcal{N} = 15$. Les résidus des méthodes de suivi $RGB + P_3$, $L_2 + P_3$ et $c_1c_2c_3 + P_3$ sont représentés sur la figure 9.3c. Tout comme pour la séquence 2, la méthode $c_1c_2c_3 + P_3$ obtient de plus faibles résidus de convergence, ce qui prouve sa très bonne précision. Notons que les attributs $A_1A_2A_3$ assurent ici une bonne robustesse du suivi. Leur apport est donc plus significatif dans le cas où des changements d'intensité d'éclairage sont provoqués, que lorsque seules des spécularités apparaissent (voir les résultats obtenus sur les séquences 1 et 2).

Séquence 4

La séquence d'images de la figure 9.4a représente une scène composée de différents objets, soumise à plusieurs éclairages (lumière du jour, tubes fluorescents au plafond, éclairage halogène direct). 15 points sont susceptibles d'être suivis au départ. À l'itération 40, l'éclairage direct est éteint brutalement, puis à l'itération 60, un tube fluorescent est à son tour éteint. Ils sont tous les deux rallumés à l'itération 80. La caméra subit un mouvement de rotation suivant l'axe optique. Le nombre de points suivis par l'approche P_3 avec normalisation L_2 ($L_2 + P_3$) ou avec invariants $c_1c_2c_3$ ($c_1c_2c_3 + P_3$) atteste là encore de leur robustesse (voir le tableau 9.4b). Les figures 9.4c et 9.4d représentent l'évolution des résidus obtenus par les méthodes les plus convaincantes (d'un côté $RGB + P_3$ et L_2 , de l'autre $L_2 + P_3$ et $c_1c_2c_3 + P_3$). Les résidus de la méthode $RGB + P_3$ atteignent des niveaux moins élevés que ceux de la norme L_2 . Le modèle photométrique s'avère mieux adapté même s'il s'agit d'une modélisation locale des variations d'illumination. Enfin, de ces quatre méthodes, $L_2 + P_3$ s'avère la plus probante, à en juger par le faible niveau de ses résidus. Notons cependant que sur cette séquence, les résidus obtenus par les différentes méthodes sont encore élevés, dans le cas où les sources d'éclairage sont éteintes (de l'itération 60 à 80). Il faut donc souligner leur faiblesse dans le cas où la saturation des couleurs est fortement réduite.

Séquence 5

La séquence 5 (voir figures 9.5a) met en jeu des objets colorés et spéculaires, qui sont immobiles. Seule la caméra bouge. Au cours de la séquence, l'ouverture de l'objectif est modifiée de sorte de faire varier l'intensité globale de l'image. 86 points sont sélectionnés au départ mais 16 d'entre eux sont perdus par occultation (voir tableau 9.5b). Néanmoins, l'approche $c_1c_2c_3 + P_3$ fournit globalement une meilleure robustesse du suivi, puisque moins de points sont perdus, et ce quel que soit \mathcal{N} . Les résidus des approches $c_1c_2c_3$ et L_2 par calcul de ϵ_1 (sans utilisation d'un modèle local d'illumination P_3) sont montrés sur la figure 9.5c. Ceux des méthodes RGB , $c_1c_2c_3$ et $L_2 + P_3$ font l'objet de la figure 9.5d. Là encore, l'association d'une correction d'illumination globale (par le biais de $c_1c_2c_3$ ou de la norme L_2) avec un modèle photométrique local (P_3), améliore la précision du suivi.

Ces différentes expérimentations ont montré l'intérêt de la couleur pour assurer le suivi de points robuste mais il s'avère important de comparer également les temps de calcul des différentes approches.

9.5.2.3 Temps de calcul

Le temps de calcul d'une méthode de suivi peut être un critère primordial dans certaines tâches de vision. Or il s'agit de la principale limitation du suivi couleur. Pour $\mathcal{N} = 11$ par exemple, la méthode KLT dans les images de luminance (calcul du mouvement à l'aide de (9.1)) prend environ 3 ms par point, tandis que la même approche en couleur prend environ 7 ms. Enfin l'approche de calcul par le biais de l'équation (9.7), prend environ 15 ms pour les images en couleur, contre seulement 5 ms pour la même approche dans les images de luminance. Précisons que ces valeurs ont été obtenues avec un processeur AMD Athlon, 1,8 GHz, avec une RAM de 512 Mo.

9.6 Conclusions

Dans ce chapitre, nous nous sommes focalisés sur le suivi de points dans les images couleur. En effet, la couleur permet généralement de sélectionner dans l'image un plus grand nombre de points en comparaison aux images de luminance.

La robustesse vis-à-vis des variations d'illumination a été améliorée de trois manières différentes : dans un premier temps, par le biais d'une modélisation photométrique, ensuite à l'aide d'invariants couleur et pour finir en exploitant conjointement ces deux approches. Les séquences d'images utilisées sont volontairement complexes (la variation de composantes tolérée est sévère compte tenu des variations d'illumination provoquées), dans le seul but de comparer les différentes approches entre elles.

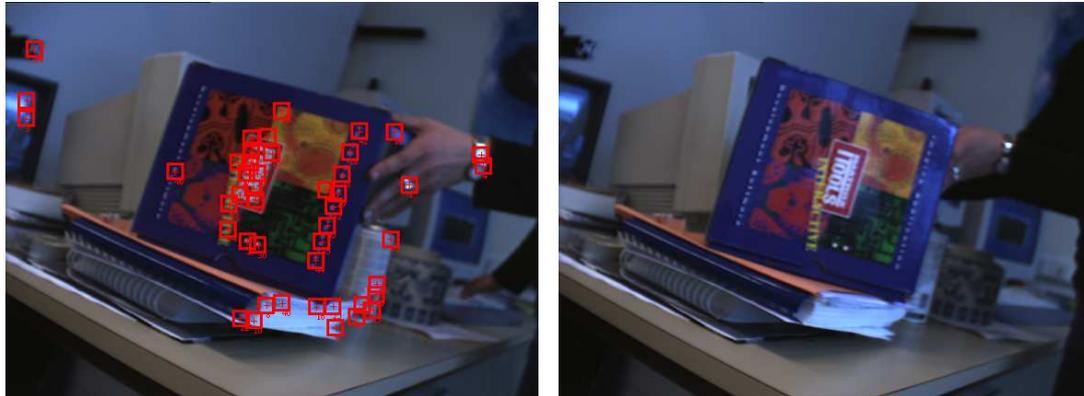
Quelques conclusions sont à souligner. Premièrement, les invariants développés pour prendre en compte les spécularités ($A_1A_2A_3$) ne s'avèrent pas toujours adaptés dans le cadre du suivi de points. D'une part, ils sont bruités dans le cas où les couleurs sont faiblement saturées

($R \simeq G \simeq B$). D'autre part, comme l'atteste l'annexe D, ils sont susceptibles de réduire la séparabilité entre les couleurs.

Il s'avère préférable d'utiliser les invariants L_2 ou $c_1c_2c_3$, valables pour des matériaux diélectriques inhomogènes et opaques seulement, et de modéliser les variations de spécularité par le modèle photométrique proposé.

Notons également que le calcul des invariants nécessite que les couleurs soient suffisamment saturées ; dans le cas contraire, il est préférable d'utiliser le modèle photométrique sans invariant ($rgb + P_3$).

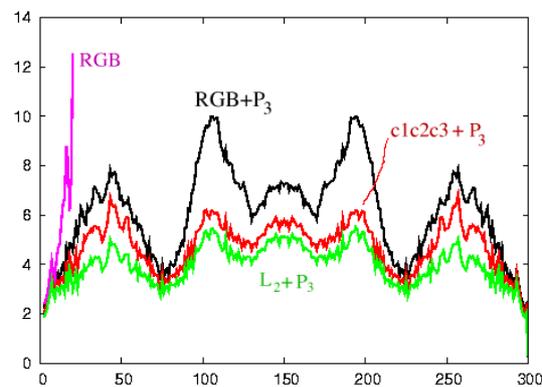
Contrairement à ce qui est mis en œuvre dans ce chapitre, la détection des points devrait être dédiée à la méthode de suivi considérée. Or, de manière à comparer les techniques de suivi entre elles, nous avons choisi de sélectionner les mêmes points pour chaque méthode.



(a)

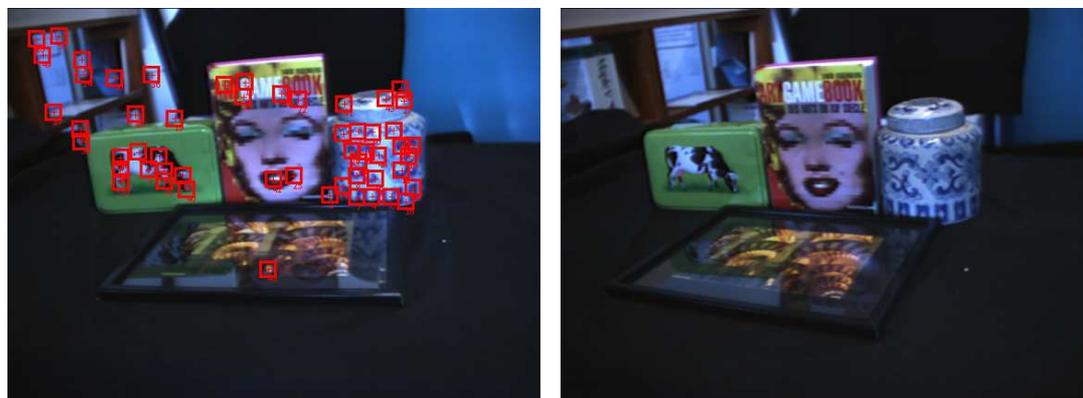
Attributs	Méthode ϵ_1 (eq. 9.1) <i>Comp</i>		Méthode ϵ_2 (eq. 9.7) <i>Comp + P₃</i>	
	$\mathcal{N}=11$	$\mathcal{N}=15$	$\mathcal{N}=11$	$\mathcal{N}=15$
C	0	0	2	9
RGB	0	0	19	25
norme L_1	0	2	3	3
norme L_2	4	5	7	18
$c_1c_2c_3$	3	3	13	20
$A_1A_2A_3$	5	0	-	-

(b)



(c)

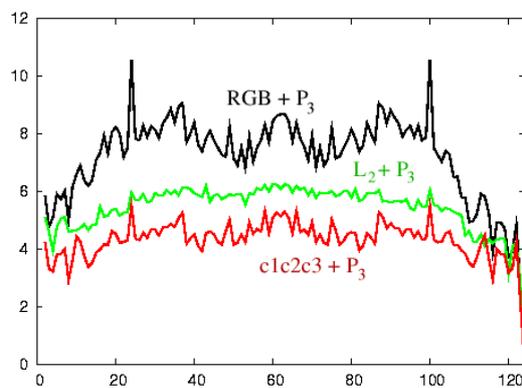
FIG. 9.2 – (a) Deux images de la séquence. (b) Nombre de points suivis (sur 36). (c) Résidus obtenus avec les méthodes RGB , $RGB + P_3$, $L_2 + P_3$ et $c_1c_2c_3 + P_3$, pour $\mathcal{N} = 15$.



(a)

Attributs	Méthode ϵ_1 (eq. (9.1)) <i>Comp</i>		Méthode ϵ_2 (eq. 9.7)) <i>Comp + P₃</i>	
	$\mathcal{N}=11$	$\mathcal{N}=15$	$\mathcal{N}=11$	$\mathcal{N}=15$
NdG	0	0	12	11
RGB	3	3	20	22
norme L_1	21	20	13	21
norme L_2	29	30	23	33
$c_1c_2c_3$	27	20	23	28
$A_1A_2A_3$	20	29	-	-

(b)



(c)

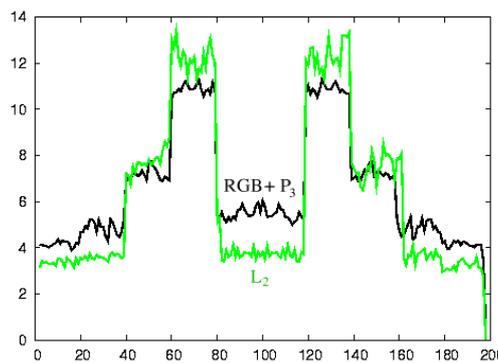
FIG. 9.3 – Séquence 3. (a) Deux images de la séquence. (b) Nombre de points suivis (sur 39). (c) Évolution des résidus pour les techniques P_3 utilisant les composantes RGB, l'invariant L_2 et l'invariant $c_1c_2c_3$.



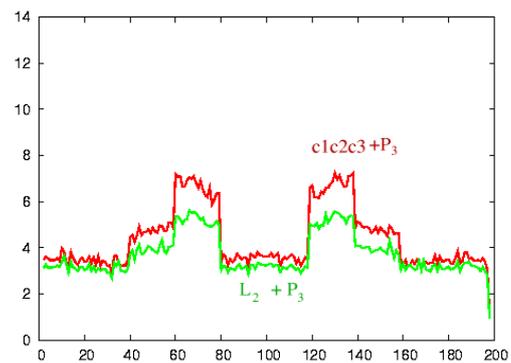
(a)

Attributs	Méthode ϵ_1 (eq. (9.1)) <i>Comp</i>		Méthode ϵ_2 (eq. (9.7)) <i>Comp + P₃</i>	
	$\mathcal{N}=11$	$\mathcal{N}=15$	$\mathcal{N}=11$	$\mathcal{N}=15$
C	0	0	0	0
RGB	0	0	3	4
norme L_1	0	0	0	3
norme L_2	2	4	5	8
$c_1c_2c_3$	3	3	3	8
$A_1A_2A_3$	3	3	-	-

(b)



(c)



(d)

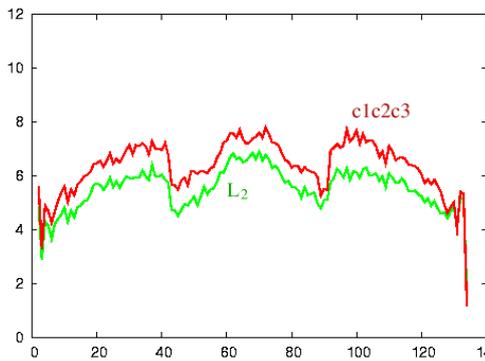
FIG. 9.4 – Séquence 4. (a) Quatre images de la séquence. (b) Nombre de points suivis (sur 15). (c) Évolution des résidus pour les techniques P_3 utilisant les composantes RGB et l'invariant L_2 . (d) Évolution des résidus des techniques $c_1c_2c_3 + P_3$ et $L_2 + P_3$ avec utilisation d'un modèle photométrique local.



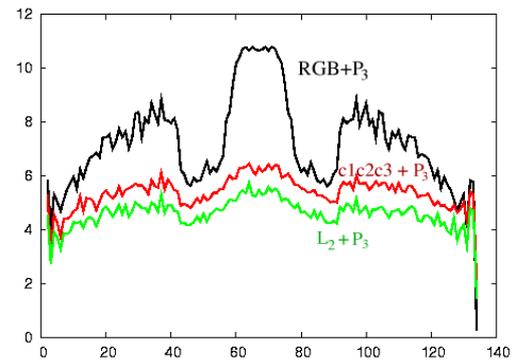
(a)

Attributs	Méthode ϵ_1 (eq. (9.1)) <i>Comp</i>		Méthode ϵ_2 (eq. 9.7)) <i>Comp + P₃</i>	
	$\mathcal{N}=11$	$\mathcal{N}=15$	$\mathcal{N}=11$	$\mathcal{N}=15$
C	0	0	7	14
RGB	0	0	9	20
norme L_1	6	5	3	8
norme L_2	3	20	16	26
$c_1c_2c_3$	14	15	8	21
$A_1A_2A_3$	9	10	-	-

(b)



(c)



(d)

FIG. 9.5 – Séquence 5. (a) Deux images de la séquence. (b) Nombre de points suivis (sur 70). (c) Résidus obtenus avec L_2 et $c_1c_2c_3$. (d) Résidus de convergence obtenus avec $rgb + P_3$, $L_2 + P_3$ et $c_1c_2c_3 + P_3$.

CONCLUSION

Cette partie du manuscrit s'est intéressée à la robustesse de suivi de primitives et du suivi de motifs vis-à-vis des variations d'illumination.

Premièrement, en se basant sur les modèles de réflexion, nous avons mis en exergue les différentes modélisations possibles de variations d'illumination produites entre deux images d'une même séquence. À partir de ces formalismes, nous avons montré sur quelles hypothèses implicites se basent les modèles photométriques locaux exploités classiquement dans les tâches de mise en correspondance et de suivi. Ces approximations supposent que les changements d'illumination soient constants en tout point de la zone de l'image considérée. Or, puisqu'elles dépendent de la géométrie de la scène (les angles d'éclairage et d'observation mesurés à partir de la normale au point considéré) mais également de la rugosité du matériau, les variations photométriques sont susceptibles d'être non constantes dans la zone d'intérêt de l'image. Nous avons alors proposé deux modèles plus appropriés à ces changements. Le premier est exploitable très localement dans l'image dans le sens où, lorsque des variations d'éclairage sont provoquées, il suppose qu'une partie de la réflectance diffuse puisse être approximée par un polynôme de degré 1. Par contre, il correspond à une modélisation très intéressante des variations de spécularités. Le second modèle prend en compte des variations plus globales des changements photométriques et s'avère valide pour de plus larges motifs dans l'image.

Ces deux modèles ont été exploités dans le cadre d'une approche de suivi différentiel. Des simulations et des résultats expérimentaux ont permis de comparer les approches proposées avec les techniques de suivi classiques. Leur pertinence a été démontrée, puisqu'elles assurent une meilleure robustesse du suivi en compensant de manière plus complète les variations d'illumination provoquées au cours des différentes séquences d'images.

Néanmoins, ces modèles se basent sur une approximation locale. À notre connaissance, il n'existe pas de modélisation globale (valable en chaque point de l'image) des variations de luminance. Par contre, des attributs invariants peuvent être définis dans les images couleur. La plupart d'entre eux se fondent toutefois sur des hypothèses quant aux matériaux mis en jeu (matériaux diélectriques inhomogènes et opaques ou matériaux conducteurs) ou encore sur la couleur de l'illuminant utilisé (illuminant blanc ou coloré). Notons que la littérature ne mentionne pas d'attributs valides pour des matériaux conducteurs éclairés par un illuminant non blanc.

Certains de ces invariants couleur ont été exploités dans le cadre du suivi de points d'intérêt. D'autre part, l'approche développée pour les images de luminance a été étendue au suivi de points d'intérêt dans les images couleur. Comme on s'y attendait, les résultats expérimentaux obtenus ont montré la bonne robustesse du modèle local dans le cas de spéularités. Par contre, dans le cas où des variations d'illumination sont considérées, les invariants couleur s'avèrent plus efficaces. Nous concilions enfin les atouts des deux types d'approches dans une approche exploitant conjointement un invariant couleur et un modèle photométrique local. En définitive, cette technique améliore les résultats du suivi quel que soit le type de variations d'illumination mis en jeu.

CONCLUSION ET PERSPECTIVES

La reconnaissance de caractères, l'asservissement visuel ou le suivi d'objets nécessitent la mise en œuvre de traitements de plus bas niveau, tels que l'extraction et le suivi de primitives. De la robustesse de ces algorithmes dépend indéniablement la fiabilité de l'application située en aval. Or, puisqu'ils se fondent généralement sur l'analyse de l'information de luminance, ces traitements peuvent se heurter aux problèmes de variations d'apparence de la scène considérée. Dans ce document, nous nous sommes concentrés sur deux problématiques liées à ces aspects. D'une part, nous avons travaillé sur la robustesse de la segmentation d'un marquage à l'encre vis-à-vis de la variation de la concentration d'encre et de la variabilité des couleurs de fond. D'autre part, nous nous sommes focalisés sur la robustesse du suivi de points et du suivi de motifs plus larges vis-à-vis des variations de spécularités et d'éclairage.



La compréhension des phénomènes d'interaction lumière-matière et l'utilisation de la couleur ont permis de proposer des solutions à ces problématiques.

Tout d'abord, après une étude de la transparence de l'encre, nous avons défini des attributs invariants vis-à-vis de l'opacité d'un marquage à l'encre et vis-à-vis de la couleur du support sur lequel il est imprimé. Ces attributs ont été exploités dans le cadre de la segmentation d'un marquage, dont la quantité d'encre mais également le support d'impression peuvent varier d'une image à l'autre. Ils autorisent une bonne adaptabilité des traitements vis-à-vis des variations de transparence de l'encre d'une image à l'autre et répondent de manière satisfaisante au problème de la variabilité des couleurs de fond. D'autre part, un atout important de cette approche réside dans la seule prise en compte de la couleur du marquage, autorisant alors des formes et des textures différentes. Une application de contrôle de la traçabilité des produits porcins, où un numéro d'abattage imprimé à l'encre sur l'animal doit être détecté, a montré toute la pertinence de ces attributs. Cet algorithme peut s'adapter à différentes couleurs d'encre et pourrait donc être utilisé dans différents abattoirs.

Ensuite, l'analyse des modèles de réflexion les plus usités en vision par ordinateur a permis de définir sur quelles hypothèses se fondent les modèles locaux de variations d'illumination,

largement exploités dans le cadre de la mise en correspondance et du suivi de points. Ce problème est en effet récurrent dans les applications de vision par ordinateur. De la même manière, les hypothèses sur lesquelles se basent les approches de suivi de points rapportées dans la littérature ont pu être explicitement décrites. Dans tous les cas, elles stipulent que les changements d'illumination sont constants, voire négligeables, au voisinage du point à suivre. Or, comme nous l'avons montré, cela n'est pas toujours réaliste, surtout si la surface considérée n'est pas plane, ou encore lorsque l'objet présente des spécularités. Nous avons présenté deux modèles photométriques permettant de compenser efficacement les variations de spécularité et les changements d'éclairage. Ils ne supposent pas que les changements d'illumination sont constants mais qu'ils varient doucement autour du point à suivre. En outre, nous avons étudié la validité de ces modèles en fonction de la géométrie de la scène et du type de surface observée. Ces modèles peuvent devenir insuffisamment précis dans des cas limites, à savoir lorsque l'objet est extrêmement spéculaire (par exemple un matériau conducteur pur, tel que le métal poli), lorsque la surface observée est discontinue ou à forte courbure, ou bien encore lorsque la source d'éclairage est très proche de la surface. Le premier des modèles n'est valable que très localement dans l'image, et s'avère tout à fait approprié au suivi de points. Le second modèle, qui compense de manière plus globale les variations d'illumination, permet d'assurer le suivi de motifs plus larges. Des résultats expérimentaux ont montré leur pertinence comparativement aux méthodes de suivi existantes, que ce soit en terme de robustesse ou de justesse de la modélisation. En outre, les temps de calcul ne sont pas accrus.

Enfin, le problème de la robustesse du suivi de points dans les images en couleur a été abordé. Tout d'abord, nous avons appliqué à ce contexte la modélisation locale proposée pour les images de luminance. D'autre part, l'utilisation d'invariants couleur a été mise en œuvre. Ces attributs compensent les variations d'illumination de manière globale (en chaque point de l'image). Dans le cas des occurrences de spécularités, nous avons montré que l'approche basée sur la modélisation locale s'avère plus pertinente qu'en utilisant des invariants couleur. Par contre, dans le cas où des variations d'intensité d'éclairage sont également provoquées, les invariants s'avèrent plus appropriés. Ce constat nous a amené à allier les avantages de ces deux types d'approches. Nous avons alors exploité conjointement l'invariance globale et le modèle local dans l'approche de suivi. Cette technique assure une très bonne robustesse et une précision du suivi, quel que soit le type de variation photométrique provoqué.



Ces travaux concernent des tâches de bas niveau pouvant être intégrées dans d'autres types d'applications.

Les invariants couleur dédiés à l'encre ont été exploités dans une tâche de segmentation pour la reconnaissance de caractères [Roué, 2004]. Dans ce cadre d'étude, l'amélioration de la segmentation a permis de faciliter et d'améliorer la tâche de reconnaissance. À notre sens, ces attributs pourraient également être mis à profit pour la restauration d'images de documents, où les encres utilisées montreraient des hétérogénéités au niveau de la quantité de l'encre.

Concernant les approches de suivi proposées, elles pourraient s'avérer tout à fait appropriées dans le cadre de l'asservissement visuel, où il s'agit par exemple d'effectuer un positionnement par rapport à un objet. Il serait intéressant d'intégrer nos travaux dans ce cadre d'étude. Nous avons déjà évoqué l'utilisation de ces approches pour le calcul de mouvement dans un contexte d'asservissement visuel [Alhaj, 2004]. En outre, la technique de suivi de points proposée peut permettre d'améliorer le positionnement de la caméra en face d'un objet [Collewet et Chaumette, 2002].

Ces travaux pourraient également s'intégrer dans les applications de suivi d'objets 3D, où les variations de spéularités peuvent apparaître. Même si des approches de type mixte, alliant à la fois les contours et la texture s'avèrent une solution intéressante à ce cadre d'étude [Masson *et al.*, 2004, Pressigout et Marchand, 2005], ces méthodes peuvent être améliorées en compensant les variations d'illumination.

D'autre part, les modèles proposés peuvent s'avérer particulièrement intéressants pour améliorer la robustesse des approches d'indexation [Gros, 2000] ou d'appariement d'images [Schmid et Mohr, 1997, Montesinos *et al.*, 2000]. En particulier, la stéréovision nécessite la mise en correspondance entre les primitives issues de deux images d'une même scène acquises selon deux angles de vue différents. Or l'importance des spéularités perçues dépend notamment de la position de l'observateur par rapport à la surface observée. À cet égard, les modèles d'illumination proposés pourraient améliorer la précision de la mise en correspondance. Dans le domaine du codage vidéo [Chaumont *et al.*, 2004, Cammas *et al.*, 2005], les variations d'illumination provoquées au cours de la séquence doivent pouvoir être codées localement et réduire ainsi les taux de compression.



Nous avons montré l'apport des modèles de réflexion et de la couleur dans les contextes d'extraction et de suivi de primitives. Toutefois, ces travaux peuvent être améliorés et poursuivis. Nous recensons ci-après les perspectives les plus intéressantes, tout d'abord dans le cas de la segmentation couleur puis dans le cas du suivi de primitives.

D'après l'étude du chapitre 3, bon nombre d'algorithmes de segmentation génériques, applicables à tout type d'images, nécessitent le réglage d'un grand nombre de paramètres ou de seuils. Il s'agit également d'une limitation de notre approche, développée dans le chapitre 4, où les paramètres sont fixés par apprentissage, de manière supervisée. Il pourrait ainsi être intéressant de mettre en œuvre une technique d'apprentissage automatique, non supervisée, de manière à fixer des paramètres et des seuils plus optimaux. Cela nécessiterait d'une part la création d'une base d'images de référence, segmentées « manuellement », d'autre part la définition de critères de choix des paramètres.

Par ailleurs, la forme du marquage à détecter est supposée inconnue et ne s'effectue qu'à l'aide d'attributs couleur. Or, dès lors qu'une base d'apprentissage n'est pas disponible, le traitement devient plus ardu, et les risques d'ambiguïtés peuvent intervenir dans la détection. Dans ce cas de figure, nous pensons que des connaissances *a priori* autres que celles de la couleur seraient nécessaires, par exemple au niveau de la forme ou de la texture du marquage

à détecter. L'algorithme de segmentation pourrait alors se baser sur la fusion éclairée de ces informations (invariants couleur dédiés à l'encre, forme, texture) à l'aide d'opérateurs de fusion appropriés [Bloch, 2003]. Une autre solution serait de sélectionner dans l'image plusieurs germes de scores d'appartenance élevés, de faire croître les régions associées à ces germes, puis de prendre finalement la décision en fonction de la forme des régions obtenues.

Enfin, la segmentation du marquage à l'encre est effectuée dans des conditions idéales d'éclairage (un éclairage diffus). Néanmoins, dans un contexte industriel, il peut s'avérer difficile de mettre en œuvre un tel système. En effet, l'encombrement de l'éclairage notamment peut constituer une réelle difficulté. Il serait alors intéressant d'étudier la faisabilité de la segmentation dans le cas d'images altérées par l'apparition d'ombres ou bien encore de spécularités. Dans ce cas de figure, l'utilisation des invariants couleur (photométriques) pourrait permettre de corriger l'apparition des ombres et de la réflexion spéculaire. Par contre, il reste encore à étudier la validité des invariants dédiés à l'encre sur ce type d'images.

Dans le cadre du suivi de primitives, des perspectives peuvent également être envisagées à court terme. Détaillons les plus pertinentes d'entre elles.

Tout d'abord, les points détectés dans l'image peuvent correspondre à des discontinuités de réflectance ou de profondeur. Or, nous avons vu que tous les modèles photométriques locaux stipulent que la région de la scène observée ne montre pas de fortes discontinuités de profondeur sur la fenêtre d'intérêt considérée. De même, dans le cas de grandes fenêtres, les points peuvent être situés à des profondeurs différentes et ne possèdent pas forcément le même mouvement, ce qui peut nuire au suivi. Dans ce cas de figure, il serait intéressant de ne sélectionner que la région pertinente de la fenêtre d'intérêt, c'est-à-dire la zone dont les points montrent la meilleure adéquation par rapport aux modèles de mouvement et d'illumination considérés. Les points aberrants, conduisant à une erreur de modélisation, seraient alors rejetés. Des approches statistiques robustes [Huber, 1981] pourraient être exploitées dans ce sens.

D'autre part, les invariants couleur permettent d'améliorer le suivi. La comparaison des différentes approches de suivi proposées a nécessité la sélection de points identiques pour toutes les techniques. Puisque les invariants couleur sont bruités dans le cas de faibles saturations des couleurs, une méthode de sélection des points prenant en compte la saturation des couleurs a été exploitée à cet effet. Nous pensons ainsi que la robustesse des approches proposées doit être améliorée par une détection de points dédiée à la technique de suivi employée.

Dans les zones de l'image où les couleurs sont faiblement saturées, le suivi couleur n'apporte pas forcément une meilleure robustesse par rapport au suivi dans les images de luminance. En effet, les invariants couleur peuvent alors s'avérer bruités dans ce cas. Par contre, ces attributs sont généralement extrêmement pertinents dans le cas d'images aux couleurs vives. D'autre part, si de fortes variations d'illumination sont provoquées dans la séquence d'images, un point de couleur fortement saturée ne le reste pas forcément tout au long de la séquence. La mise en œuvre d'un algorithme compétitif entre l'approche de suivi couleur (utilisant les invariants et la modélisation locale) et la méthode de suivi dans les images de luminance (avec la modélisation photométrique locale) permettrait d'améliorer la robustesse. À notre sens, cela nécessite la définition d'un critère de comparaison de la « pertinence » de chacune des approches, qui se

baserait sur la valeur de saturation des couleurs et sur les résidus de convergence. La définition d'un tel critère constitue une perspective de recherche en soi.



Outre les améliorations possibles des travaux présentés dans ce document, il est possible d'envisager des travaux de recherche à plus long terme.

Les modèles photométriques ont été exploités dans le but de compenser les variations de spécularités. Dans une optique différente, il serait intéressant de les utiliser dans un contexte de vision active où il s'agirait de réagir aux configurations d'observation de la scène pour lesquelles elles apparaissent. En effet, les paramètres d'illumination calculés nous renseignent sur l'évolution des changements de spécularités au cours de la séquence d'image. Ils peuvent s'exprimer analytiquement en fonction de l'angle d'observation, de la normale au point considéré, de la direction d'éclairage, des caractéristiques du matériau. À partir de connaissances *a priori* sur le contexte, même grossières, (par exemple la position de la source, la forme de l'objet, la position d'observation), il pourrait être intéressant d'estimer leurs valeurs en fonction du déplacement de la caméra. Cette dernière pourrait alors éviter les configurations conduisant à l'apparition de fortes spécularités et permettrait de rendre plus robuste la tâche à réaliser.

Enfin, la plupart des techniques d'asservissement visuel repose sur l'existence de primitives géométriques dans l'image, voire la connaissance d'un modèle 3D de l'objet observé. Or, dans le cas d'environnements complexes ou inconnus, ces informations ne sont pas toujours disponibles. Notamment, la connaissance du modèle 3D d'un objet est une contrainte forte. L'utilisation d'informations photométriques est une réponse appropriée à ce cas de figure. Néanmoins, peu de travaux traitent de ce problème, du fait de la difficulté d'obtenir une relation entre la variation d'intensité mesurée dans l'image et le déplacement effectué par la caméra. Toutefois, dans [Deguchi et Noguchi, 1996], une transformation de Karhunen-Loeve a été proposée. Dans ce type d'approche, la prise en compte des variations d'illumination au cours du mouvement de la caméra devrait permettre d'assurer une meilleure précision du positionnement de la caméra.

MESURE DES PARAMÈTRES DE RÉFLEXION DU JAMBON

Mesurer la radiance d'un objet en plusieurs positions de la source d'éclairage permet de connaître la manière dont il renvoie l'énergie lumineuse. Cette radiance peut ensuite être modélisée (voir la section 1.2). La fonction de modélisation fournit des informations quant aux spécularités, à leurs positions, à leur importance. Dans le cadre de notre étude de traçabilité, nous nous intéressons à des produits porcins. L'objectif de cette étude consiste notamment à évaluer l'importance de la réflexion spéculaire du jambon. D'autre part, nous cherchons à déterminer les directions de spécularités privilégiées de cet objet. Celles-ci peuvent être utiles pour éviter leur occurrence dans la phase d'acquisition d'images.

Ces mesures ont été réalisées en utilisant un système existant, développé dans le cadre d'une application textile, où les objets traités sont de très faible épaisseur. Ce système est illustré sur la figure A.1. Son adaptation à notre étude implique certaines corrections.

1) Correction des défauts liés au système de mesure

Du fait de l'épaisseur de la pièce de porc, la surface de mesure n'est pas située au centre du demi-cercle représenté sur la figure A.1, ce qui provoque des imprécisions sur les angles de mesure et une variation complexe de l'angle solide suivant l'angle d'incidence. En outre, nous avons tenu à conserver la pièce de porc entière (et non un fin morceau de couenne) afin que la radiance mesurée prenne en compte la constitution interne de la pièce (présence de chair, de sang, aspect légèrement translucide de la couenne). Un deuxième défaut est dû à la forme de l'objet lui-même, puisqu'il est difficile d'obtenir une surface parfaitement plane, dont la normale coïncide parfaitement avec l'origine du repère. Ensuite, l'hétérogénéité de la pigmentation de la couenne, que ce soit sur un même jambon ou d'un jambon à l'autre, rend nécessaire la multiplicité des mesures. Nous comparons ici les mesures effectuées sur deux jambons. Elles sont accomplies sur une surface circulaire de 3 cm de diamètre environ. Notons que l'ampleur de la

spécularité peut dépendre partiellement de l'humidité de la couenne issue de la condensation, qui s'avère extrêmement variable et donc très difficile à contrôler.

La radiance \mathcal{L} est mesurée avec un spectroradiomètre et une image de la surface est acquise à chaque position d'éclairage considérée. En outre, aucun éclairage ambiant n'est pris en compte. La radiance est mesurée pour chaque position θ_i de la lumière selon les angles zénithaux, mais seulement pour quelques positions d'observation θ_r . On considère que l'origine des angles coïncide avec la normale à la surface d'intérêt (voir figure A.1), et que ces angles sont signés. Étant donné que la surface de mesure n'est pas située au centre du système, la distance entre la source et la surface, et donc l'angle solide d'éclairage $d\Omega$ varie en fonction de θ_i . Par conséquent, l'intensité énergétique reçue par la surface varie également. Par contre, étant donné que l'appareil de mesure n'est pas déplacé, l'angle solide de mesure n'est pas modifié. Il faut donc normaliser les valeurs de radiance obtenues.

Soit $d(\theta_i)$ la distance entre la source et la surface et R le rayon du demi-cercle. On note également a l'épaisseur de la pièce de porc, avec $a \ll R$, et \mathcal{A} l'aire de la surface considérée. Ensuite, l'angle solide d'éclairage est donné par :

$$\Omega = \frac{\mathcal{A} \cos \theta_i}{d(\theta_i)} \quad (\text{A.1})$$

L'intensité $I(\theta_i)$ reçue par la surface à l'angle θ_i est alors définie comme :

$$I(\theta_i) = \frac{\Phi_i}{\Omega} \quad (\text{A.2})$$

Considérons que le flux énergétique Φ_i produit par l'éclairage est constant au cours du temps, quelle que soit la position de la source. Dans le cas où l'éclairage coïncide avec la normale à la surface, pour $\theta_i = 0$, la distance source/surface est donnée par $d(0) = R - a$. Pour les autres positions d'éclairage, la distance entre la source et l'éclairage vaut $d(\theta_i) \simeq R - a \sin \theta_i$, dans le cas où $a \ll R$, par trigonométrie. Afin d'obtenir en tout angle θ_i la même intensité énergétique d'éclairage qu'en $\theta_0 = 0$, il faut multiplier l'intensité en θ_i par le coefficient \mathcal{C}_E suivant :

$$\mathcal{C}_E = \frac{(R - a)^2}{(R - a \sin \theta_i)^2} \cos \theta_i \quad (\text{A.3})$$

2) Mesure et calcul des paramètres de réflexion

Les figures A.5a, A.5b et A.5c montrent un exemple de mesure de radiance ainsi que les approximations obtenues. Sur la première d'entre elles figure la composante diffuse. Dans un premier temps, nous modélisons la composante de réflexion diffuse du jambon.

• Composante diffuse

En fonction de l'allure des mesures de réflectance (figure A.5), il est assez aisé de déterminer approximativement les positions d'angles correspondant à une absence ou à une présence

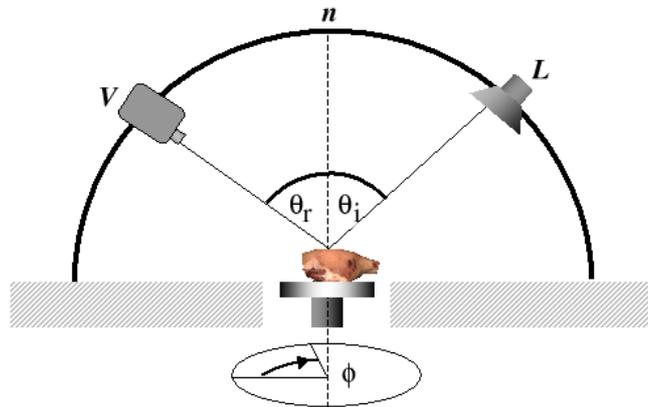


FIG. A.1 – Système de mesure utilisé. V est la direction d'observation, L la direction d'éclairage, n la normale à la surface considérée.



FIG. A.2 – Différentes images de la couenne, pour différentes positions d'éclairage, la direction d'observation étant fixe, $\theta_r = -45$. (a) $\theta_i = 65$. (b) $\theta_i = 40$. (c) $\theta_i = 0$. (d) $\theta_i = -40$.

de composante spéculaire. Néanmoins, il s'avère plus précis de se référer au modèle dichromatique [Shafer, 1985]. En effet, comme le décrit l'annexe B, une classe de couleur en forme de "T" atteste de la présence de spécularités. Ainsi, les positions zénithales pour lesquelles il n'y a aucune composante spéculaire sont recherchées, puis la composante diffuse est modélisée à partir de ces mesures. La figure A.2 montre quatre images d'une même zone de couenne acquises en quatre positions d'éclairage.

La composante diffuse de la radiance obtenue, appelée $\mathcal{L}^d(\theta_i)$ ne peut pas exactement être modélisée par un modèle lambertien. Nous admettons que cet artefact est lié à la configuration du système, notamment à l'imprécision des angles (due à l'épaisseur de la pièce de porc), ainsi qu'au fait que la source n'est en pratique pas tout à fait ponctuelle. Nous parvenons à la modéliser par une fonction cosinus munie d'un exposant :

$$\mathcal{L}^d(\theta_i) = A_0 \cos^n(\theta_i) \quad (\text{A.4})$$

Les angles de mesure ne correspondent pas exactement aux angles d'incidence, là encore à cause du décalage de la pièce de porc par rapport au centre du système. Plus précisément, les angles mesurés sont plus petits que les angles d'incidence effectifs. Afin de corriger le défaut

de positionnement du jambon, dont la normale à la surface considérée ne coïncide pas exactement avec l'origine du repère (voir la figure A.1), nous ajoutons au modèle un terme angulaire correctif ρ_0 :

$$\mathcal{L}^d(\theta_i) = A_0 \cos^n(\theta_i - \rho_0) \quad (\text{A.5})$$

où ρ_0 , A_0 et n correspondent à des constantes calculées par interpolation au moyen des moindres carrés. Le tableau A.1 contient les valeurs obtenues sur les différentes pièces de porc. Nous n'avons pas spécifié les valeurs de ρ_0 , qui ne sont pas propres au matériau, mais dépendent de la configuration du système de mesure. Il est intéressant de souligner la variation de K_0 d'un terme à l'autre, qui s'explique par la différence d'albedo d'un jambon à l'autre.

• Composante spéculaire

Une fois que la composante de réflexion diffuse est modélisée, la composante spéculaire est approximée par un modèle de Phong

$$\mathcal{L}^s = K_1 \cos^{n_1}(\rho + \rho_1) \quad (\text{A.6})$$

ρ_1 est un terme correctif de la direction de spécularité, ici encore introduit pour prendre en compte les approximations de la mesure. Les valeurs calculées sont regroupées dans le tableau A.2 avec des valeurs de rugosité n_1 comprises entre 60 et 90 environ, et des valeurs de K_1 avoisinant 0.2. Ce modèle est amélioré en lui rajoutant un second lobe spéculaire doté de termes n_2 , ρ_2 et K_2 . Ce lobe, de rugosité comprise entre 300 et 400, est à mettre en relation avec le modèle de Nayar, présenté dans la section 1.2.2.5, où la modélisation d'un lobe spéculaire est effectuée en additionnant un lobe spéculaire et un pic de réflexion presque idéalement spéculaire.

La grande surprise de cette mesure est la présence d'un second lobe de réflexion spéculaire. Sur la figure A.4, il est centré sur l'origine du repère ($\theta_i \simeq 0$), lorsque $\theta_r = -\pi/4$. L'explication est à rechercher soit au niveau de l'aspect de surface de la couenne soit au niveau de la structure interne du matériau, la couenne étant légèrement translucide. La réponse à cette question nécessiterait une étude plus poussée sur le matériau, ce qui ne fait pas l'objet de nos travaux. Si l'on privilégie l'hypothèse liée à la surface du matériau, une façon de modéliser la reflectance est alors de considérer la surface de la couenne comme une combinaison de deux types de surfaces d'orientations différentes, où le second lobe de réflexion est modélisé comme dans (A.7) en y ajoutant un terme θ_n correspondant à l'angle entre les deux surfaces :

$$\mathcal{L}^s = K_3 \cos^{n_3}(\rho + \theta_n + \rho_3) \quad (\text{A.7})$$

Le terme ρ garde la même signification que précédemment, et ρ_3 correspond à un terme correctif au même titre que ρ_1 et ρ_2 . D'après les mesures, θ_n serait à peu près égal à 45. De même que pour le lobe spéculaire décrit par (A.6), la modélisation du second lobe de réflexion peut être améliorée en lui adjoignant un pic spéculaire de paramètres n_4 et ρ_4 . Le tableau A.3 regroupe les valeurs obtenues.

• Étude de l'anisotropisme.

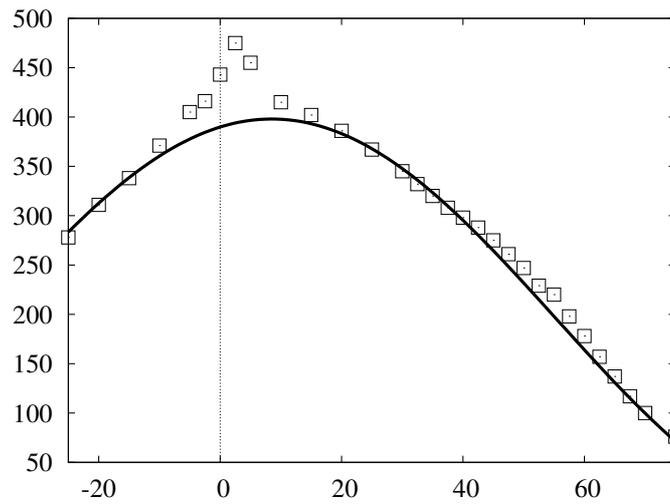


FIG. A.3 – Mesure de la radiance (points carrés) et modélisation de la composante diffuse de réflexion (en trait plein)

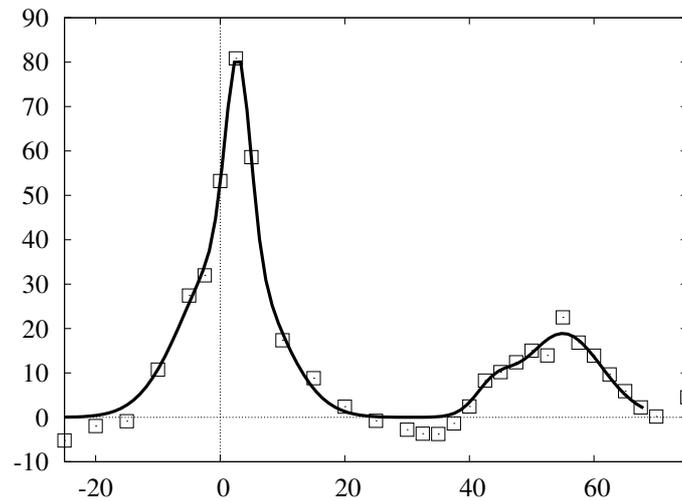


FIG. A.4 – Modélisation de la composante de réflexion spéculaire (en trait plein)

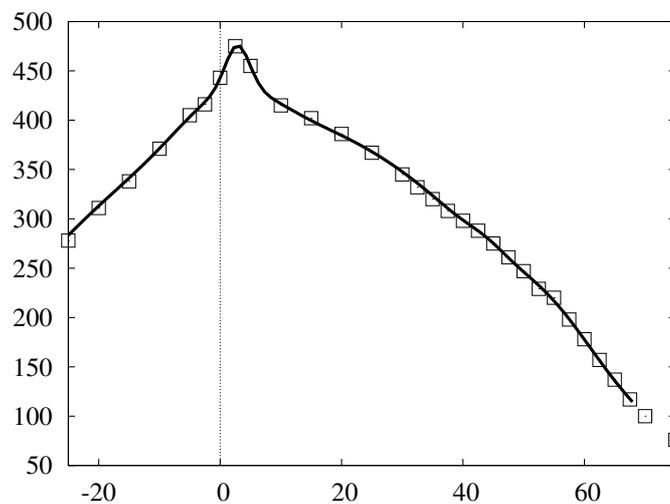


FIG. A.5 – Modélisation par le modèle de Phong pour un angle d'observation $\theta_r = 45$.

Lorsqu'un matériau est isotrope, la radiance ne dépend pas de l'azimut ϕ_i . Le caractère anisotrope d'un matériau dépend généralement de la qualité de surface, dans le cas par exemple où elle est striée ou composée de fibres alignées. La mesure de ce caractère anisotrope s'effectue par rotation de la pièce de porc d'un angle ϕ autour de la normale \mathbf{n} , le spectrophotomètre restant fixe. Ces mesures ont également été modélisées par un modèle de Phong.

L'origine des angles ϕ a été choisie de manière arbitraire, globalement parallèle à l'axe principal de jambon. La figure A.6 montre les résultats de mesure de la reflectance du jambon suivant les angles azimuthaux. On remarque sur cet exemple que la fonction de radiance comprend deux lobes principaux, à 50 et 275 degrés. La surface du jambon est donc anisotrope. En effet, en se référant aux différentes images de la figure A.2, on note la rugosité inhérente de la couenne mais également une aspérité due à la présence de poils. Si l'on modélise ces mesures par une fonction cosinus, comme pour le modèle de Phong, on obtient la formulation qui suit :

$$\begin{cases} \mathcal{L}(\phi) = b_0 + b_1 \cos^{n_5}(\phi - \rho_5) & \text{pour } \phi \in [0, 150[\\ \mathcal{L}(\phi) = b_0 + b_2 \cos^{n_6}(\phi - \rho_6) & \text{pour } \phi \in [150, 360] \end{cases} \quad (\text{A.8})$$

où b_0, b_1, b_2, n_5, n_6 et ρ_5, ρ_6 sont les constantes à déterminer. D'après nos expériences, $n_5 \simeq 6.5$ et $n_6 \simeq 1.5$. Ensuite, ρ_5 et ρ_6 sont à peu près égaux à $5\pi/18$ (50) et $55\pi/36$ (275) respectivement. Les termes b_0, b_1 et b_2 varient d'un jambon à l'autre, comme nous pouvons le voir dans le tableau A.4.

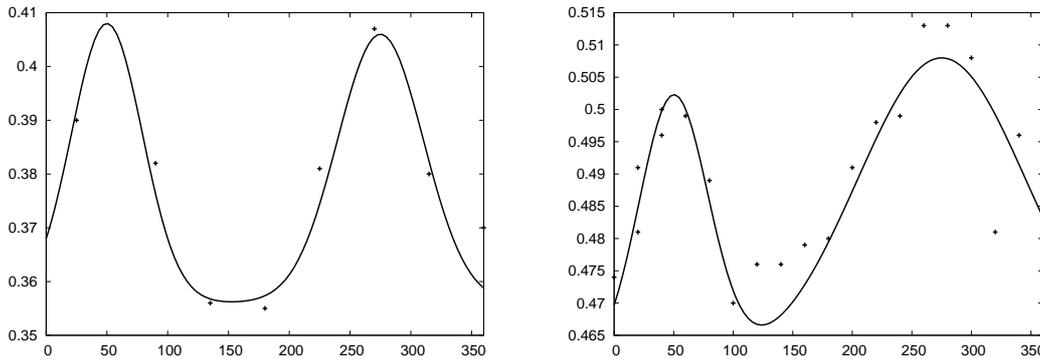


FIG. A.6 – Modélisation de la reflectance en fonction de l'azimut par le modèle de Torrance. (a) Jambon A. (b) Jambon B.

3) Conclusion

Dans cette annexe, nous avons présenté un exemple de mesure de la radiance du jambon et déterminé ses paramètres de réflexion. Ceci nous a permis de comparer les résultats de deux jambons de manière quantitative.

Cette étude nous apporte plusieurs renseignements. Premièrement, le jambon n'est pas un objet lambertien puisqu'il montre même deux directions privilégiées de réflexion. L'importance de ces lobes spéculaires varie d'un jambon à l'autre mais également en fonction de l'humidité de la pièce de porc. De même, les paramètres de réflexion montrent de fortes variations d'un

TAB. A.1 – Paramètres du modèle lambertien obtenus avec deux jambons A et B.

Paramètres	A	B
K_0	0.398	0.349
n	1.87	1.63

TAB. A.2 – Paramètres du modèle de Phong pour le lobe spéculaire.

Jambon	Paramètres					
	K_1	K_1	ρ_1	ρ_2	n_1	n_2
A	0.019	0.0059	10	-1	85.7	367
B	0.0262	0.010	10	-1	62	300

TAB. A.3 – Paramètres du modèle de Phong pour le second lobe de réflexion.

Jambon	Paramètres					
	K_3	K_4	ρ_3	ρ_4	n_3	n_4
A	0.0383	0.0439	1.4	2.9	63	841
B	0.0507	0.0119	1.5	15.7	98.6	396

TAB. A.4 – Paramètres du modèle de Phong en fonction de l'azimut.

Zone	Paramètres		
	b_0	b_1	b_2
Jambon A	0.355	0.051	0.054
Jambon B	0.335	0.026	0.037

jambon à l'autre, notamment au niveau de l'*albedo*. Enfin, nous avons pu montrer qu'il ne s'agit pas d'un matériau isotrope, du fait de la structure de sa surface.

Fort de ce constat, il paraît nécessaire d'adapter le système d'éclairage à l'objet considéré afin d'obtenir des images dépourvues de spécularités. L'utilisation d'un système d'éclairage diffus s'avère la meilleure solution ; néanmoins, un tel type de configuration peut être difficile à utiliser dans un cadre industriel.

Les spécularités peuvent être évitées en positionnant l'éclairage au plus près de la caméra. En effet, le jambon montre une réflexion exclusivement diffuse dans ce cas de figure (voir la figure A.5 pour $\theta_i < -10$).

LE MODÈLE DICHROMATIQUE

Soit P un point de la scène et p sa projection dans l'image. L'expression du modèle dichromatique de radiance au point P , décrit auparavant dans la section 1.2.2.6, s'écrit :

$$\mathcal{L}(\lambda, P) = l_s(\lambda, P)m_s(P) + l_b(\lambda, P)m_b(P) \quad (\text{B.1})$$

$l_s(\lambda, P)$ est assimilé à la répartition spectrale du terme de réflexion spéculaire, qui correspond à celle de l'illuminant $l_s(\lambda, P) = \mathcal{E}(\lambda, P)$. Quant à $l_b(\lambda, P)$, il correspond à la répartition spectrale de la radiance due à la réflexion diffuse : $l_b(\lambda, P) = \mathcal{R}_b(\lambda, P)\mathcal{E}(\lambda, P)$, avec $\mathcal{R}_b(\lambda, P)$ la réflectance du matériau associée à la réflexion diffuse. Ce modèle est essentiellement qualitatif, dans la mesure où il ne précise pas de modèle pour les fonctions $m_b(p)$ et $m_s(p)$. Par intégration du stimulus par le capteur tri-CCD, de sensibilités $S_i(\lambda)$ ($i = R, G, B$), on fait intervenir explicitement le vecteur couleur associé à la réflexion diffuse $\mathbf{C}_b(p) = (C_b^R, C_b^G, C_b^B)^T$ et le vecteur couleur associé à la réflexion spéculaire $\mathbf{C}_s(p) = (C_s^R, C_s^G, C_s^B)^T$ au pixel p :

$$C_b^i(p) = \int_{\lambda} S_i(\lambda)\mathcal{E}(\lambda, P)\mathcal{R}_b(\lambda, P)d\lambda, \quad C_s^i(p) = \int_{\lambda} S_i(\lambda)\mathcal{E}(\lambda, P)d\lambda \quad (\text{B.2})$$

Ainsi, le modèle dichromatique de Shafer, exprimé dans l'espace RGB , devient :

$$\mathbf{C}(p) = m_b(p)\mathbf{C}_b(p) + m_s(p)\mathbf{C}_s(p) \quad (\text{B.3})$$

Avec $m_b(p) = m_b(P)$. D'après le modèle dichromatique de Shafer (B.3), l'ensemble des couleurs d'un matériau est représenté dans l'espace couleur par un plan défini par les couleurs $\mathbf{C}_s(p)$ et $\mathbf{C}_b(p)$, comme l'illustre la figure B.1. Les couleurs d'un matériau lambertien, de réflexion purement diffuse, sont situées sur un segment C_b . D'après [Klinker *et al.*, 1988, Klinker *et al.*, 1990], les couleurs du matériau spéculaire sont localisées avec une plus grande précision encore, sur un *cluster* en forme de T (voir figure B.1b). Dans les cas idéaux où l'objet est de couleur uniforme et de surface très lisse [Nayar *et al.*, 1997], la barre verticale du T part de l'origine $RGB = (0, 0, 0)$ et pointe vers la couleur de la composante lambertienne du matériau et le pied du T correspond exactement à la couleur de l'illuminant.

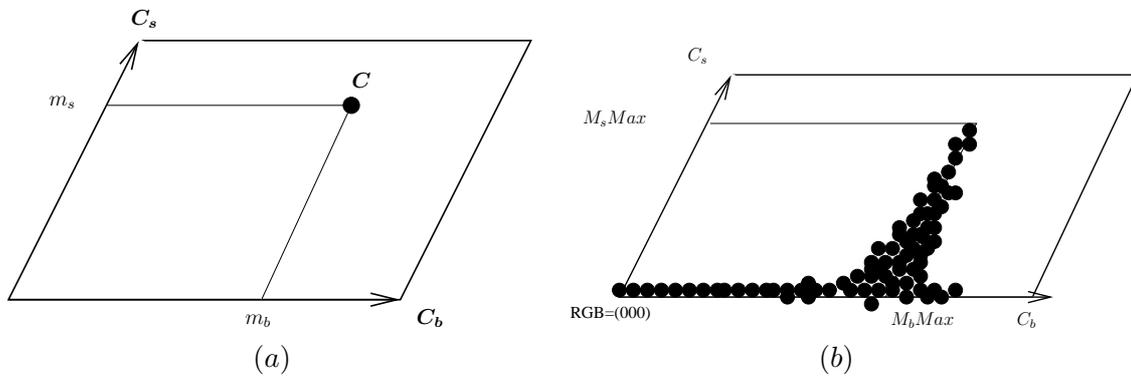


FIG. B.1 – Réflexion d'une couleur d'après le modèle dichromatique de Shafer. (a) Représentation des couleurs dans un plan (parallélogramme de Shafer). (b) Représentation des couleur par une classe de couleurs en forme de T.

ROBUSTESSE DES MÉTHODES DE SUIVI COMPENSANT LES VARIATIONS D'ILLUMINATION

Dans cette annexe, nous étudions tout d'abord l'influence d'un pixel bruité sur le calcul de la normalisation photométrique effectuée dans la technique de suivi décrite dans le paragraphe 7.3.2 page 155. Dans un second temps, nous montrons que l'approche de suivi basée sur une modélisation affine des variations d'illumination (voir la section 7.3.1) est moins bien conditionnée que l'approche de suivi de points proposée dans la section 7.4.

C.1 Erreur provoquée par un pixel bruité sur la normalisation photométrique

Soit $v(p)$ la normalisation photométrique, telle que $v(p) = \frac{f(p) - \mu_f}{\sigma_f}$, où μ_f et σ_f sont respectivement la moyenne et l'écart-type des luminances sur la fenêtre d'intérêt \mathcal{W} , de taille $\mathcal{N} \times \mathcal{N}$. Si l'on considère qu'un pixel p_0 de \mathcal{W} de valeur $f(p_0)$ est bruité, la moyenne μ_f s'exprime :

$$\mu_f = \frac{f(p_0)}{\mathcal{N}} + \frac{(\mathcal{N} - 1)\mu}{\mathcal{N}}$$

μ étant la moyenne des pixels de \mathcal{W} qui ne sont pas entachés de bruit. On exprime l'écart-type σ_f sur le même principe.

Ainsi, en admettant que toute erreur σ_v^2 sur $v(p)$ est provoquée exclusivement par la valeur bruitée f_0 , une erreur de mesure de $v(p)$ se calcule à partir de l'erreur provoquée sur $f(p_0)$, appelée $\sigma_{f_0}^2$:

$$\sigma_v(p)^2 = \left(\frac{\partial v(p)}{\partial f(p_0)} \right)^2 \sigma_{f(p_0)}^2 \quad (\text{C.1})$$

Tous calculs faits, la dérivée au carré de v par rapport à f_0 est exprimée par un rapport A/B :

$$\left(\frac{\partial v(p)}{\partial f(p_0)} \right)^2 = \frac{A}{B} \quad (\text{C.2})$$

où A et B sont exprimés de la manière suivante :

$$\begin{aligned} A &= [(6\mu^2 - (3f(p) + f(p_0))\mu + f(p)f(p_0))\mathcal{N}^3 + (-12\mu^2 + S_{f_2} + (5f(p) - f(p_0))\mu - f(p)f(p_0))\mathcal{N}^2 + \\ &\quad (8\mu^2 + 2(f(p_0) - f(p))\mu)\mathcal{N} - 2\mu^2 - 2f(p_0)\mu]^2 \mathcal{N}^2 \\ B &= [3\mu^2\mathcal{N}^3 + (f^2(p_0) - 4\mu^2 - 6f(p_0)\mu + S_{f_2})\mathcal{N}^2 + (\mu^2 - f^2(p_0) + 8f(p_0)\mu)\mathcal{N} - 4f(p_0)\mu] \end{aligned}$$

où le terme S_{f_2} est la somme des luminances (sauf $f(p_0)$) au carré :

$$S_{f_2} = \sum_{p \in \mathcal{W}, p \neq p_0} f^2(p) \quad (\text{C.3})$$

Comparons le rapport de (C.1) par rapport à l'unité, ou ce qui revient au même la différence entre le numérateur et le dénominateur par rapport à 0. Dans le cas où il est supérieur à l'unité, un bruit provoqué sur p_0 influence énormément la valeur de $v(p)$. Dans le cas contraire, le bruit est atténué par la normalisation. Afin de simplifier les calculs, considérons que $f(p) = \mu$, c'est-à-dire que l'image est uniforme. Dans ce cas, la différence entre le numérateur et le dénominateur de la relation (C.1) s'exprime par :

$$A - B = -4\mu^2\mathcal{N}^3 + (8\mu^2 + 4f(p_0)\mu)\mathcal{N}^2 + (-10f(p_0)\mu + f(p_0)^2 - 3\mu^2)\mathcal{N} + 4f(p_0)\mu \quad (\text{C.4})$$

Cette valeur est toujours négative, comme le montre la figure C.1. Notamment, dans le cas où $f(p_0) = \mu$, la valeur (C.4) est donnée par :

$$A - B = -4\mu^2(\mathcal{N} - 1)^3 \quad (\text{C.5})$$

qui est négative pour toute taille de fenêtre $\mathcal{N} > 1$. En outre, plus la taille de la fenêtre \mathcal{N} est grande plus la différence $A - B$ est négative et moins le bruit apparaissant sur un pixel influe sur la valeur de normalisation photométrique.

C.2 Conditionnement de la méthode de suivi de Soatto

Les méthodes de suivi de points différentielles passent par la minimisation d'une fonction de coût, qui nécessite l'inversion d'une matrice, $\sum_{p \in \mathcal{W}} \mathbf{V} \mathbf{V}^T$ (voir la relation (7.7) page 152), avec dans le cas du suivi classique un vecteur \mathbf{V} exprimé par (7.8).

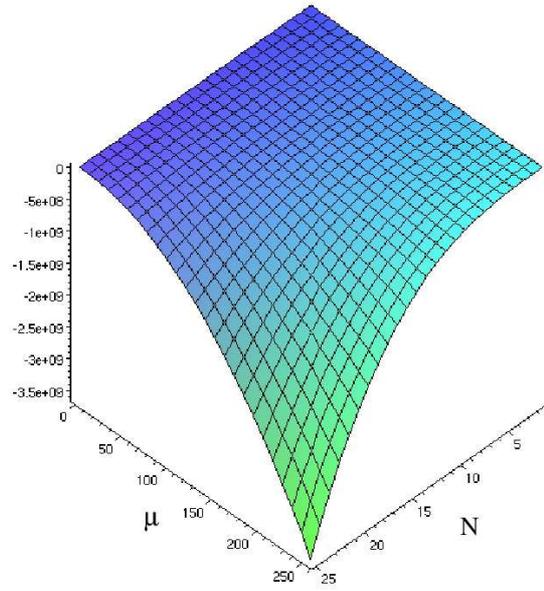


FIG. C.1 – Variations de la valeur (C.4) en fonction de μ , variant entre 0 et 255, et de \mathcal{N} variant entre 0 et 25, pour $f(p_0) = 255$. Le rapport A-B est toujours négatif : une erreur sur $f(p_0)$ est atténuée par la normalisation photométrique.

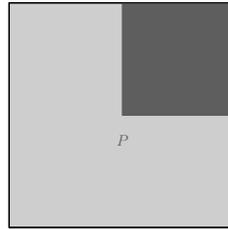


FIG. C.2 – Niveaux d'intensité d'une fenêtre d'intérêt représentant un coin formant un angle de 90 .

Une méthode de suivi peut être précise si cette matrice est suffisamment bien conditionnée. Or, dans le cas de la méthode compensant le modèle affine d'illumination [Jin *et al.*, 2001], la matrice mise en jeu $\mathbf{M} = \sum_{p \in \mathcal{W}} \mathbf{V}_s \mathbf{V}_s^T$ dépend du contenu de l'image f , comme le montre l'expression du vecteur \mathbf{V}_s , donnée par (7.13). Afin d'étudier le conditionnement de cette matrice, nous considérons un point situé sur un coin idéal, formant un angle à 90, comme le montre la figure C.2, avec une valeur de luminance nulle pour les pixels associés au coin et de 250 pour les autres pixels. Il s'agit d'un motif parmi les plus susceptibles d'être extraits par la méthode de détection de Harris [Harris et Stephens, 1988].

Soit le nombre de conditionnement $K(\mathbf{M}) = \|\mathbf{M}\|_{\infty} \|\mathbf{M}^{-1}\|_{\infty}$. Dans le cas d'une fenêtre d'intérêt de taille $\mathcal{N} = 9$, la matrice mise en jeu dans l'approche exploitant un modèle photométrique affine obtient une valeur de conditionnement égale à 100601. L'approche que nous proposons ne fait pas intervenir les valeurs de luminance dans le calcul de la matrice. Le condi-

tionnement obtenu pour $\mathcal{N} = 9$ vaut alors 41386. La méthode proposée s'avère donc mieux conditionnée.

COMPARAISON DES INVARIANTS

Dans cette annexe, nous comparons quelques espaces d'attributs invariants parmi ceux décrits dans le chapitre 8. Nous considérons les espaces normalisés par les normes L_1 et L_2 ainsi que les composantes $c_1c_2c_3$, dédiés aux matériaux diélectriques, inhomogènes et opaques. Nous étudions également quelques attributs invariants valables à la fois pour les matériaux conducteurs et les matériaux diélectriques inhomogènes et opaques, tels que la teinte H , issue de l'espace HSV , les composantes $l_1l_2l_3$, ainsi que l'espace d'attributs proposé $A_1A_2A_3$. Soit f et g deux images exprimées dans un espace couleur ou un espace d'invariants, et acquises dans des conditions d'éclairage différentes. Le critère d'invariance \mathcal{CI}_i d'un attribut C_i (pour $i = 1..3$) est fourni par le calcul du pourcentage de variation des invariants d'une image à l'autre :

$$\mathcal{CI}_i = (100/\mathcal{N}) \left(\sum_{\mathcal{N}} |f_i - g_i| \right) / |f_i| \text{ pour } i = 1..3 \quad (\text{D.1})$$

où \mathcal{N} est la taille de l'image, en pixels. Plus la valeur de ces critères est faible, plus l'attribut est considéré invariant.

Considérons différentes scènes simples présentant des variations d'éclairage ou de spécularité. Dans tous les cas, l'illuminant est considéré blanc (une balance des blancs a été effectuée).

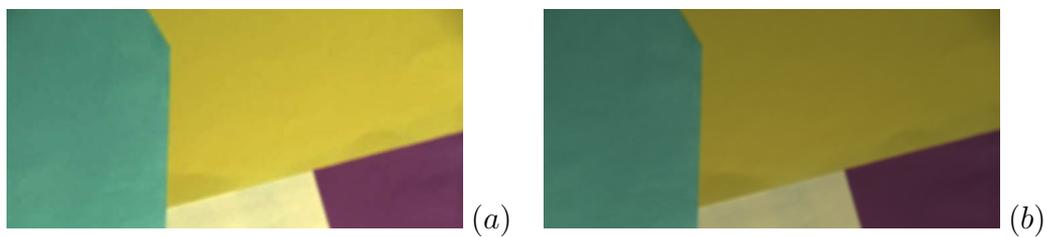
Les images D.1a et D.1b représentent des surfaces colorées, correspondant à des feuilles de papier qui sont éclairées sous deux positions différentes d'éclairage. Notons que l'on ne distingue pas de spécularité, puisqu'il s'agit de matériaux diélectriques opaques. D'après le tableau D.1c regroupant les valeurs \mathcal{CI}_i obtenues à partir de ces images, tous les invariants comparés obtiennent de bien meilleures valeurs d'invariance (critères \mathcal{CI}_i plus faibles) que les composantes RGB . Alors que les images RGB montrent de fortes variations d'illumination, de 27% à 30%, les valeurs associées aux attributs invariants ne dépassent pas 5%. Soulignons en particulier les bonnes performances des attributs définis pour prendre en compte les variations de spécularité, $l_1l_2l_3$ et $A_1A_2A_3$. Ils présentent les meilleures invariances parmi les attributs testés, puisque leurs critères sont dans tous les cas inférieurs à 3. Par ailleurs, les normes L_1 et L_2

tout comme les attributs $c_1c_2c_3$, détiennent des résultats très comparables. Ils se basent effectivement sur les mêmes hypothèses quant au type de matériau considéré. La figure D.1d montre quelques images d'invariants calculés à partir des images D.1a et D.1b. Dans tous les cas, les variations d'illumination et les ombres provoquées par les plis du papier ne sont quasiment plus perceptibles. Néanmoins, les invariants adaptés aux matériaux spéculaires (la teinte H et les attributs $l_1l_2l_3$ et $A_1A_2A_3$), en plus de réduire les variations d'illumination, réduisent également la discrimination entre certaines couleurs. Les images obtenues avec ces attributs ne montrent plus aucune différence entre les deux zones jaunes présentes sur les images D.1a et D.1b. Ce constat peut s'avérer pénalisant dans le cadre du suivi de primitives. Cela est d'autant plus vrai dans le cas de la teinte.

Les images D.2a et D.2b correspondent à des surfaces colorées, respectivement lambertiennes et spéculaires, qui ont été acquises sur des objets réels. Dans les deux cas, chaque rangée d'images est associée à des conditions différentes d'éclairage. Dans le premier cas (D.2a), des variations d'intensité sont provoquées. Dans le second cas (D.2b), des changements de direction d'éclairage induisent des variations de spécularité. Les valeurs \mathcal{CT}_i obtenues sont regroupées dans le tableau D.2c. Notons là encore que tous les attributs apportent une bonne invariance vis-à-vis des variations d'illumination considérées, puisque les valeurs de critères sont dans tous les cas plus faibles que pour les images RGB . Cependant, les attributs H , $l_1l_2l_3$ et $A_1A_2A_3$ n'obtiennent pas les résultats escomptés, notamment dans le cas de variations de spécularité, dans le sens où ils n'apparaissent pas plus invariants que les attributs normalisés et $c_1c_2c_3$. Or ils sont théoriquement appropriés aux matériaux spéculaires. Ce constat peut s'expliquer par leur comportement bruité dans le cas de couleurs peu saturées. Or, paradoxalement, les zones spéculaires s'avèrent généralement peu saturées.

Intéressons-nous maintenant au cas où changements d'éclairage et occurrences de spécularités sont simultanément visibles. Les images D.3a et D.3b présentent des surfaces colorées spéculaires éclairées selon deux positions différentes d'éclairage. Alors que les images en RGB montrent jusqu'à 80 % de variations photométriques, les attributs invariants obtiennent des valeurs inférieures à 10 %. Notons ici les très bonnes performances des attributs $l_1l_2l_3$ et $A_1A_2A_3$ pour lesquels les valeurs \mathcal{CT}_i sont plus faibles que ceux calculés à partir des invariants dédiés aux matériaux lambertiens (normes L_1 et L_2 , $c_1c_2c_3$).

Conclusion. Dans cette annexe, quelques invariants couleur ont été comparés. Certains d'entre eux sont en théorie dédiés aux matériaux diélectriques inhomogènes et opaques (norme L_1 et L_2 , attributs $c_1c_2c_3$) tandis que les autres doivent pouvoir s'appliquer aux matériaux conducteurs (la teinte H , les attributs $l_1l_2l_3$). Toutes ces composantes apportent effectivement de très bonnes propriétés d'invariance photométrique comparativement aux coordonnées classiques RGB , quelle que soit la cause du changement d'illumination (des variations de spécularité ou des changements d'éclairage). Les composantes H , $l_1l_2l_3$ et $A_1A_2A_3$ peuvent répondre à l'invariance vis-à-vis des occurrences de spécularités tant que les couleurs de l'objet considérées sont suffisamment saturées. Dans le cas contraire, ces attributs s'avèrent généralement bruités. Par ailleurs, nous avons vu à travers un exemple (voir la figure D.1d), qu'ils réduisaient la séparabilité entre certaines couleurs proches, ce qui n'est pas le cas des invariants dédiés aux matériaux lambertiens.



Attributs	Composante		
	1	2	3
RGB	29,0	29,8	27,7
L_1	1,6	1,6	4,0
L_2	1,5	1,3	4,3
$c_1c_2c_3$	1,7	1,5	4,4
H	3	-	-
$l_1l_2l_3$	0,59	2,1	2
$A_1A_2A_3$	0,06	2	1,9

(c)

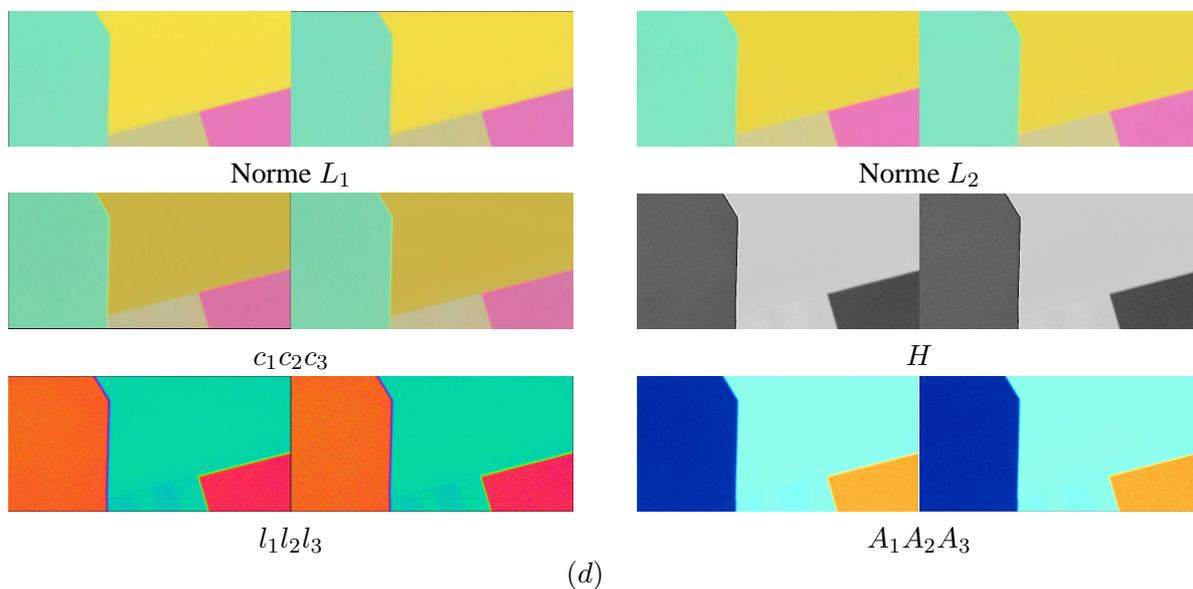


FIG. D.1 – Images de test. (a) Image initiale. (b) Image après variation d'illumination. (c) Valeurs CI_i obtenues sur les images (a) et (b). (d) Images d'invariants associées aux figures (a) et (b).

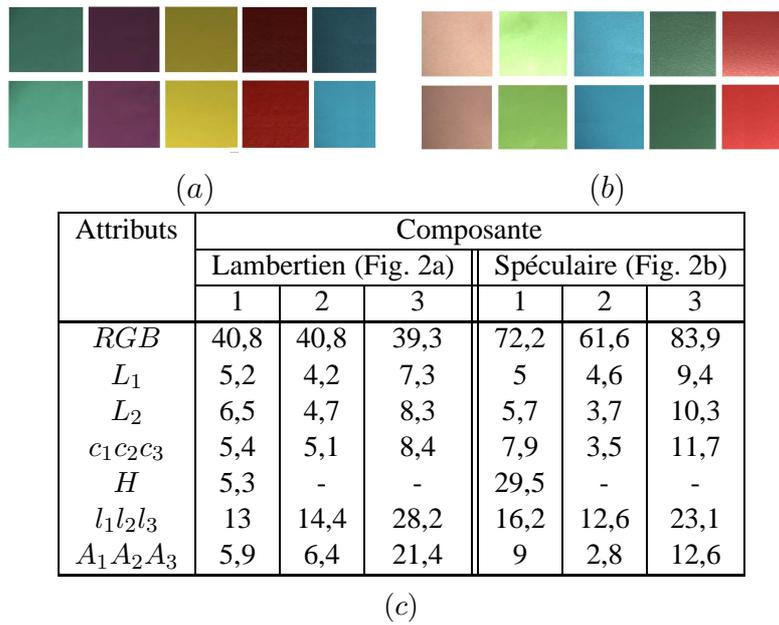


FIG. D.2 – Test de l'invariance. (a) Image de surfaces lambertiennes. (b) : Images de surfaces spéculaires. (c) : Critères CI_i obtenus sur les images (a) (première colonne) et (b) (seconde colonne).

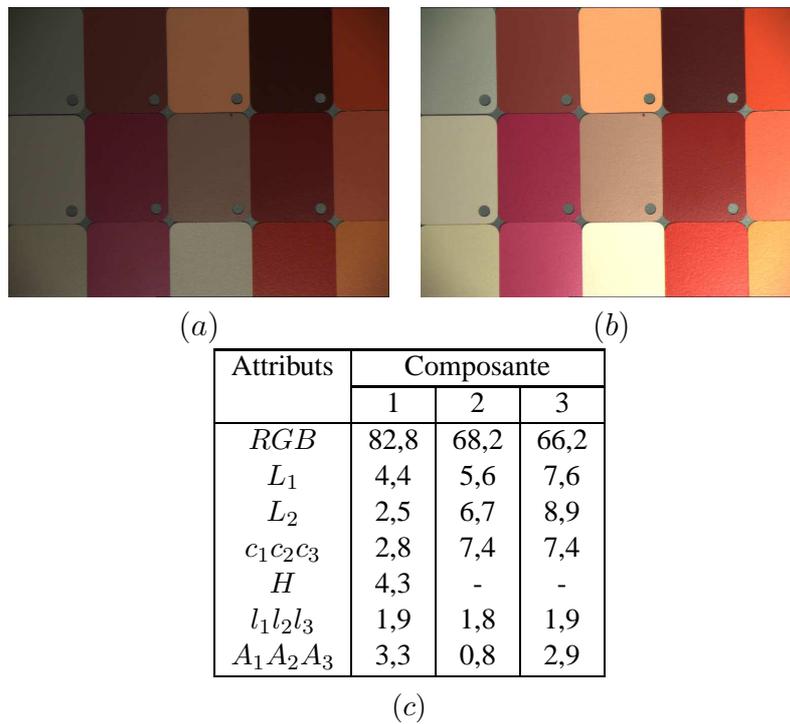


FIG. D.3 – (a) Image initiale. (b) Image après variation d'illumination. (c) Valeurs CI_i calculées entre les images (a) et (b).

GLOSSAIRE

Les définitions qui suivent ont été inspirées pour la plupart de plusieurs documents [Hunter et Harold, 1987, Desvignes, 1991, Joly, 1999, Trouvé, 1991, AFE, 1991].

Absorption : phénomène selon lequel une radiation électromagnétique est convertie en un autre type de radiation (une onde lumineuse est convertie en chaleur).

Achromatique : se dit d'une couleur sans *chrominance** (comme le blanc, le gris, le noir).

Albedo : rapport de l'énergie lumineuse réfléchie ou diffusée par une surface sur l'énergie lumineuse incidente sur cette même surface. L'albedo est également appelé *coefficient de réflexion*.

Apparence : aspect visuel très général qui caractérise les lumières et les objets de notre environnement : transparence, brillance, couleur, texture.

Balance des blancs : équilibrage des gains des trois canaux d'une caméra couleur.

Candela : unité d'intensité lumineuse. Un candela vaut un *lumen* par stéradian.

Chroma : pureté de la couleur, pouvant être exprimée dans les systèmes uniformes CIELUV et CIELAB.

Chromaticité : caractéristiques définissant la position d'une couleur dans le plan chromatique soit par définition de ses coordonnées chromatiques (x et y , ou a^* et b^* par exemple) soit par définition de sa *teinte* et de sa *pureté*.

Chrominance : écart colorimétrique entre deux couleurs de même luminance Y et donc situées sur le plan chromatique. Cette grandeur est souvent donnée pour la comparaison d'une couleur par rapport à une couleur de référence achromatique de même *luminosité**.

Clarté : attribut de la sensation visuelle selon lequel un corps paraît transmettre ou réfléchir par diffusion une fraction plus ou moins grande de la lumière incidente. C'est le correspondant psychosensoriel de la grandeur photométrique *facteur de luminance lumineuse*. La sensation de clarté est fortement liée au fond coloré sur lequel est posé l'échantillon observé.

Colorimétrie : étude des phénomènes en relation avec la couleur et leur perception par l'œil humain. La colorimétrie est basée sur l'effet trichrome de la sensation visuelle.

Composantes trichromatiques : quantités des trois stimuli de couleur de référence, qui, dans un système trichromatique donné, sont nécessaires pour égaliser la couleur du stimulus considéré (voir *égalisation**).

Conductivité électrique [S/m] : capacité de la matière à mettre en mouvement ses électrons sous l'action d'un champ électrique.

Constance colorimétrique : invariance d'une couleur d'un objet vis-à-vis de la couleur de l'illuminant.

Coordonnées trichromatiques (x, y, z) : rapport de chacune des trois *composantes trichromatiques* à leur somme.

Diagramme de Voronoï : structure divisant l'espace euclidien, à partir de germes d'entrée, en régions \mathcal{R}_i appelées régions de Voronoï. Une région de Voronoï \mathcal{R}_i contient tous les points les plus proches du germe s_i que de n'importe quel autre germe.

Diffuseur : élément réémettant la lumière incidente dans toutes les directions, quel que soit l'angle d'incidence. On distingue les diffuseurs par réflexion des diffuseurs par transmission (corps translucide).

Diffuseur parfait : objet blanc dont le coefficient de réflexion (ou *albedo**) est supposé unitaire (l'énergie lumineuse émise par sa surface est supposée égale à l'énergie reçue). Ses coefficients de transmission et d'absorption sont nuls.

Écart chromatique : différence visuelle perceptible entre deux couleurs.

Eclairement lumineux [lx] : rapport entre le flux lumineux et la zone illuminée. Un éclairement de 1 lux se crée lorsqu'un flux lumineux de 1 *lumen** est distribué de façon égale sur une zone de 1 mètre carré.

Égalisation : procédé visant à reproduire visuellement, à partir d'un mélange pondéré de trois faisceaux de couleurs primaires, un stimulus chromatique de référence de longueur d'onde donnée.

Facteur de luminance : rapport de la luminance du corps considéré sur la luminance d'un *diffuseur parfait** dans des conditions d'éclairement et d'observation identiques.

Flux lumineux [lm] : puissance irradiée émise par une source lumineuse et perçue par l'œil.

Fovéa : partie centrale de la tache jaune située sur la surface rétinienne, où se concentre la majorité des cônes récepteurs.

Illuminant : rayonnement dont la répartition spectrale relative d'énergie est définie dans le domaine des longueurs d'onde capables d'influencer la perception de la couleur de objets.

Intensité : l'une des trois caractéristiques permettant de définir une couleur. L'intensité correspond à la *luminosité** d'une couleur, donc à une quantité d'énergie.

Intensité lumineuse [cd] : intensité d'un rayonnement visible dans une direction particulière. L'ensemble de ces rayonnements constitue le flux lumineux.

Interreflexion : phénomène par lequel un rayon est émis sur une surface d'un objet et est réémis par cette surface vers une surface voisine.

Irradiance : terme anglo-saxon désignant l'*éclairement énergétique**.

Longueur d'onde : distance de propagation d'une onde électromagnétique pendant la période T , exprimée par le rapport de la célérité de la lumière sur la fréquence $\lambda = C/\nu$.

Longueur d'onde dominante : longueur d'onde monochromatique qui, mélangée à une couleur achromatique, restitue une impression colorée équivalente à un stimulus donné.

Lumen [L] : unité d'éclairement.

Lumière monochromatique : onde électromagnétique de fréquence unique. Elle présente une double périodicité temporelle (T) et spatiale (ν).

Luminance lumineuse [cd/m^2] : flux lumineux par unité d'angle solide, émis par une source primaire ou surface illuminée (source secondaire) dans une direction donnée. La luminance d'un objet dans une scène caractérise sa clarté par rapport au blanc de la scène.

Luminance énergétique [$\text{Wm}^{-2}\text{sr}^{-1}$] : flux énergétique par unité d'angle solide, émis par une surface illuminée.

Luminosité : perception visuelle de la luminance. Il s'agit de l'intensité de la couleur. La luminosité fait référence à la quantité de lumière qui est absorbée ou réfléchi par la zone colorée, elle fait donc appel à la sensibilité de la vision humaine aux intensités lumineuses. L'AFNOR la définit comme l'attribut de la sensation visuelle selon lequel une surface éclairée par une source lumineuse déterminée paraît émettre plus ou moins de lumière.

Lux : unité d'éclairement correspondant à un lumen par m^2 [$\text{L}\cdot\text{m}^{-2}$].

Matériau conducteur : matériau pour lequel la réflexion est uniquement surfaciale. L'onde incidente ne pénètre presque pas dans le matériau.

Matériau diélectrique : matériau pour lequel les ondes électromagnétiques, en plus d'être réfléchies, sont réfractées voire transmises. Le verre par exemple est un matériau diélectrique.

Matériau optiquement homogène : matériau constitué de particules aux propriétés optiques identiques.

Métamérisme : phénomène caractérisant des rayonnements de couleur ayant les mêmes composantes trichromatiques, bien que les répartitions spectrales soient différentes.

Onde électromagnétique : onde transversale composée d'une onde de champ électrique et d'une onde de champ magnétique.

Onde monochromatique : onde électromagnétique d'un rayonnement lumineux de longueur d'onde donnée. La couleur de cette lumière est liée à la valeur de sa fréquence.

Opacité : capacité d'un corps à s'opposer à la transmission d'un flux lumineux.

Perméabilité magnétique [H/m] : terme physique représentant la capacité de la matière à se polariser sous l'action d'un champ magnétique.

Permittivité électrique [F/m] : terme physique caractérisant la polarisation (la redistribution locale des charges liées) d'un matériau sous l'action d'un champ électrique.

Photométrie : science décrivant d'un point de vue psychophysique la luminance perçue par un capteur. Si le capteur est un œil humain, les grandeurs photométriques se déduisent des grandeurs *radiométriques* en pondérant la distribution spectrale du rayonnement par l'*efficacité lumineuse relative spectrale* de l'œil.

Planckian : un illuminant est dit planckian si son spectre correspond à celui de l'émission d'un corps noir, dont la *température de couleur** est définie.

Polarisation de la lumière : orientation du champ électrique (ou magnétique) dans le plan perpendiculaire à la propagation de l'onde.

Pureté : finesse spectrale d'une couleur donnée. Elle est définie comme le rapport de la *luminance énergétique** spectrale à la *longueur d'onde dominante**, sur la luminance énergétique spectrale totale.

Radiance : terme anglo-saxon désignant la *luminance énergétique**.

Radiométrie : branche de l'optique permettant de décrire d'un point de vue physique l'énergie émise par un rayonnement optique, non nécessairement visible.

Reflectance : rapport du flux réfléchi par l'échantillon étudié, sur le flux incident. Plus précisément, il s'agit du rapport entre la *radiance** émise par une microsurface dans une direction et l'*irradiance** reçue par cette même microsurface dans une autre direction.

Reflectance spectrale (ou facteur spectral de diffusion) : rapport, à une longueur d'onde donnée, de la luminance énergétique d'une surface sur celle d'un diffuseur parfait.

Réflexion diffuse : phénomène de dispersion de l'énergie incidente sur une surface dans toutes les directions de l'espace.

Réflexion spéculaire : phénomène qui redirige un rayon de lumière incident à la surface d'un matériau selon un angle de réflexion. Par exemple, les miroirs idéaux admettent une réflexion spéculaire pour laquelle l'angle du rayon émis est égal à l'angle du rayon incident. La part d'énergie réfléchie par réflexion spéculaire est donnée par le coefficient de Fresnel.

Réfraction : changement angulaire subi par un rayon lumineux pénétrant de façon oblique un milieu d'indice de réfraction différent.

Saturation : sensation visuelle permettant d'estimer la proportion de couleur pure dans une sensation couleur globale. Une couleur saturée (ou pure) ne contient pas de blanc. Les couleurs non saturées correspondent aux niveaux de gris. Physiquement la teinte est la proportion d'énergie émise à la longueur d'onde de tonalité, par rapport à l'ensemble de l'énergie émise par le spectre. Ainsi, une couleur très pure présentera un spectre pointu alors qu'une couleur impure, constituée d'un mélange de plusieurs longueurs d'onde, présentera un spectre étalé.

Sensibilité lumineuse spectrale : efficacité du capteur en fonction de la longueur d'onde. Pour l'œil humain, cette sensibilité, définie de manière statistique, est appelée visibilité.

Source : objet capable de produire un rayonnement à partir d'une transformation d'énergie. Le spectre de la source peut être discontinu, continu ou mixte.

Source primaire de lumière : surface ou objet émettant de la lumière produite par une transformation d'énergie.

Source secondaire de lumière : surface ou objet qui, n'émettant pas de lumière par lui-même, reçoit de la lumière et la restitue, au moins partiellement, par réflexion ou par transmission. Une source se différencie de l'illuminant dans le sens où la source est la réalisation pratique de l'illuminant.

Spectrocolorimétrie : système de mesure permettant de déterminer les *composantes trichromatiques* à partir de la mesure du facteur spectral de luminance.

Synthèse additive : méthode de création des couleurs par addition de faisceaux lumineux (généralement de couleurs rouge, vert et bleu).

Synthèse soustractive : méthode de création des couleurs par soustraction à la lumière blanche des composantes bleue, verte et rouge, par l'emploi de filtres respectivement jaune, magenta et cyan.

Teinte ou tonalité chromatique : attribut de sensation visuelle qui conduit à définir la perception comme bleue, vert, rouge, jaune, etc. Elle correspond d'un point de vue physique à la *longueur d'onde dominante** de la couleur : dans le cas d'une couleur monochromatique, il s'agit tout simplement du maximum spectral (par exemple pour le vert : 546.1 nm) mais lorsque la couleur est issue d'un mélange, la longueur d'onde dominante correspond au centre de gravité du spectre.

Température de couleur : la température de couleur d'une source lumineuse est définie en comparaison avec un corps radiateur noir rapporté sur la *courbe de Planck*. Plus la température de ce corps est élevée, plus grande est la composante bleue du spectre et plus petite la composante rouge. Une lampe à incandescence émettant de la lumière blanche chaude, par exemple a une température de couleur de 2.700 K tandis que celle d'une lampe fluorescente à lumière du jour a une température de couleur de 6.000 K [Trémeau *et al.*, 2004]. La température de couleur s'exprime en Kelvin [K].

Transmission : phénomène par lequel une onde électromagnétique traverse un objet.

Vision mésopique : vision dite « crépusculaire », intermédiaire entre la *vision photopique** et la *vision scotopique**. La sensibilité spectrale se déplace alors vers les petites longueurs d'onde donnant une sensation bleutée.

Vision photopique : vision dite « de jour » dans le cas où l'éclairement est suffisant ($3,4 \text{ cd/m}^2$). Le maximum de sensibilité est alors de 560 nm, c'est-à-dire dans la sensation colorée verte.

Vision scotopique : vision dite « de nuit » pour des éclairagements faibles ($<0,034 \text{ cd/m}^2$).

BIBLIOGRAPHIE

- [Aaron et Steven, 2005] AARON et STEVEN, M. (2005). Exemple-based photometric stereo : shape reconstruction with general varying BRDF. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264.
- [AFE, 1991] AFE (1991). *La photométrie en éclairage*. Société d'éditions LUX.
- [Agoston, 1987] AGOSTON, G. (1987). Color theory and its application in art and design. *Springer Verlag, optical Science series*, pages 66–77.
- [Alhaj, 2004] ALHAJ, A. (2004). *Apport de la vision dynamique en asservissement visuel, Thèse de doctorat*. Université de Rennes I, Cemagref Rennes.
- [Altunbasak et al., 2003] ALTUNBASAK, Y., LERSEREAU, R. et PATTI, A. (2003). A fast parametric motion estimation algorithm with illumination and lens distortion correction. *IEEE Transactions on Image Processing*, 12(4):395–408.
- [Angulo et Serra, 2003] ANGULO, J. et SERRA, J. (2003). Color segmentation by ordered mergings. *In IEEE International Conference on Image Processing*, volume 2, pages 125–128, Barcelona, Spain.
- [Arnaud et al., 2004] ARNAUD, E., MÉMIN, E. et FRIAS, B. C. (2004). Filtrage conditionnel pour la trajectographie dans des séquences d'images - application au suivi de points. *In 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, RFIA'04*, Toulouse, France.
- [Asada et Brady, 1986] ASADA, H. et BRADY, M. (1986). The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):2–14.
- [Augereau et al., 2005] AUGEREAU, B., TREMBLAIS, B. et FERNANDEZ-MALOIGNE, C. (2005). A vectorial computation of the optical flow in color image sequences. *In IS&T Color Imaging Conference*, Scottsdale, Arizona, USA.
- [Baker, 2004] BAKER, S. (2004). Lucas-kanade 20 years on : a unifying framework. *International journal of computer vision*, 56(3):221–255.
- [Ballerini, 2003] BALLERINI, L. (2003). Genetic snakes for color images segmentation. *In Lecture notes in computer vision*, page 268. Springer-Verlag.
- [Banks et Corke, 2001] BANKS, J. et CORKE, P. (2001). Quantitative evaluation of matching methods and validity measures for stereo vision. *International Journal of Robotics Research*, 20(7):400–409.

- [Barnard *et al.*, 1997] BARNARD, K., FINLAYSON, G. et FUNT, B. (1997). Color constancy for scenes with varying illumination. *Computer Vision and Image Understanding*, 65(2):311–321.
- [Basri et Jacobs, 2001] BASRI, R. et JACOBS, D. (2001). Photometric stereo with general, unknown lighting. In *IEEE Computer Vision and Pattern Recognition*, volume 2, pages 374–381.
- [Beaudet, 1978] BEAUDET, P. (1978). Rotationally invariant image operators. In *Proceedings of the 4th International Conference on Pattern Recognition*, pages 579–583.
- [Beck *et al.*, 1984] BECK, J., PRAZDNY, K. et IVRY, R. (1984). The perception of transparency with achromatic colors. *Perception and psychophysics*, 35(10):407–422.
- [Beckmann et Spizzichino, 1987] BECKMANN, P. et SPIZZICHINO, A. (1987). *The scattering of electromagnetic waves from rough surfaces*. Artech House Inc, 2 édition.
- [Berry, 1987] BERRY, D. (1987). Colour recognition using spectral signatures. *Pattern Recognition Letters*, 6:69–75.
- [Bezdeck *et al.*, 1984] BEZDECK, J., EHRLICH, R. et FULL, W. (1984). FCM : the fuzzy c-means clustering algorithm. *Computer and Geosciences*, 10(2-3):191–203.
- [Black *et al.*, 2000] BLACK, M., FLEET, D. et Y. YACOOB (2000). Robustly estimating changes in image appearance. *Computer Vision and Image Understanding*, 78:8 – 31.
- [Bloch, 1996] BLOCH, I. (1996). Information combination operators for data fusion. *IEEE Transactions on Systems, Man and Cybernetics*, 26(1):52–67.
- [Bloch, 2003] BLOCH, I. (2003). *Fusion d'informations en traitement du signal et des images*. Traitement du signal et de l'image, Hermes, Lavoisier, Paris.
- [Borsotti *et al.*, 1998] BORSOTTI, M., CAMPADELLI, P. et SCHETTINI, R. (1998). Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19(8):741–747.
- [Braddick *et al.*, 2002] BRADDICK, O., WISHART, K. et CURRAN, W. (2002). Directional performance in motion transparency. *Vision Research*, 42(10):1237–1248.
- [Brill, 1990] BRILL, M. H. (1990). Image segmentation by object color : a unifying framework and connection to color constancy. *Journal of the Optical Society of America A*, 7(10):2041–2047.
- [Brun, 2002] BRUN, L. (2002). *Traitements d'images couleurs et pyramides combinatoires*, Mémoire d'Habilitation à Diriger des Recherches. Université de Reims.
- [Brun et Tremeau, 2002] BRUN, L. et TREMEAU, A. (2002). *Color quantization*, chapitre 9, pages 589–637. Electrical and Applied Signal Processing Series. CRC Press.
- [Buenaposada et Baumela, 2002] BUENAPOSADA, J. et BAUMELA, L. (2002). Real-time tracking and estimation of plane pose. In *IEEE International Conference on Pattern Recognition*, volume 2, pages 697–700, Quebec, Canada.
- [Cammass *et al.*, 2005] CAMMNAS, N., PATEUX, S., MORIN, L. et GUEN, B. L. (2005). Estimation de mouvement par maillages et lignes de rupture pour le codage de séquences vidéo. In *20ème colloque GRETSI sur le Traitement de Signal et des Images*, Louvain-La-Neuve.
- [Campadelli *et al.*, 1997] CAMPADELLI, P., MEDICI, D. et SCHETTINI, R. (1997). Color image segmentation using HOPFIELD networks. *Image and Vision Computing*, 15(3):161–166.

- [Canny, 1986] CANNY, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.
- [Carron, 1995] CARRON, T. (1995). *Segmentation d'images couleur dans la base Teinte-Luminance-Saturation : approche numérique et symbolique*. Thèse de doctorat, Université de Savoie, Annecy, France.
- [Caselles *et al.*, 1993] CASELLES, V., CATTE, F., COLL, T. et DIBOS, F. (1993). A geometric model for active contours in image processing. *Numerische Mathematik*, 66:1–31.
- [Celenk, 1990] CELENK, M. (1990). A color clustering technique for image segmentation. *Computer Vision Graphics Image Process*, 52:145–170.
- [Chan *et al.*, 2000] CHAN, T., SANDBERG, B. Y. et VESE, L. (2000). Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11:130–141.
- [Chassery et Garbay, 1984] CHASSERY, J. M. et GARBAY, C. (1984). An iterative segmentation method based on a contextual color and shape criterion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):794–800.
- [Chaumont *et al.*, 2004] CHAUMONT, M., PATEUX, S. et NICOLAS, H. (2004). Object-based video coding using a dynamic coding approach. In *IEEE International Conference on Image Processing*, pages 1105–1108, Singapour.
- [Chen *et al.*, 1995] CHEN, C.-H., LEE, J.-S. et SUN, Y.-N. (1995). Wavelet transformation for gray-level corner detection. *Pattern Recognition*, 28(6):853–861.
- [Chen et Lu, 2002] CHEN, T. et LU, Y. (2002). Color image segmentation - an innovative approach. *Pattern Recognition*, 35(2):395–405.
- [Chen *et al.*, 2003] CHEN, T., MURPHEY, Y., KARLSEN, R. et GERHART, G. (2003). Color image segmentation in color and spatial domain. In *International Conference on Industrial, Engineering and other Applications Applied to Intelligence Systems*, pages 72–82.
- [Cheng *et al.*, 2001] CHENG, H., JIANG, X., SUN, Y. et WANG, J. (2001). Color image segmentation : advances and prospects. *Pattern Recognition*, 12(34):2259–2281.
- [Cheng *et al.*, 2002] CHENG, H., JIANG, X. et WANG, J. (2002). Color image segmentation based on homogram thresholding and region merging. *Pattern Recognition*, 35(2):373–393.
- [Cheng et Li, 2003] CHENG, H. et LI, J. (2003). Fuzzy homogeneity and scale space approach to color image segmentation. *Pattern Recognition*, 36(7):1545–1562.
- [Chien et Sung, 2000] CHIEN, S.-I. et SUNG, S.-H. (2000). Adaptive window method with sizing vectors for reliable correlation-based target tracking. *Pattern Recognition*, 33:237–249.
- [Claustres, 2003] CLAUSTRES, L. (2003). *Modélisation de la fonction de distribution de réflectance bidirectionnelle par ondelette pour le rendu physiquement réaliste*. Thèse de doctorat, Université Paul Sabatier.
- [Cocquerez et Philipp, 1995] COCQUEREZ, J. et PHILIPP, S. (1995). *Analyse d'images filtrage et segmentation*. Enseignement de la physique, Masson, Paris.
- [Colantoni, 1998] COLANTONI, P. (1998). *Contributions des structures de données à la segmentation d'images couleur. Élaboration d'un outil d'infographie textile*. Thèse de doctorat. Université Jean Monnet, St Etienne.

- [Collewet *et al.*, 2004] COLLEWET, C., ALHAJ, A. et CHAUMETTE, F. (2004). Model-free visual servoing on complex images based on 3D reconstruction. *In IEEE Int. Conf. on Robotics and Automation*, New Orleans, USA.
- [Collewet et Chaumette, 2002] COLLEWET, C. et CHAUMETTE, F. (2002). Positioning a camera with respect to planar object of unknown shape by coupling 2D visual servoing and 3D estimations. *IEEE Transactions on Robotics and Automation*, 18(3):322–333.
- [de Broglie, 1924] de BROGLIE, L. (1924). *Recherches sur la théorie des quanta*. Thèse de doctorat.
- [Deguchi et Noguchi, 1996] DEGUCHI, K. et NOGUCHI, T. (1996). Visual servoing using eigenspace method and dynamic calculation of interaction matrices. *In International Conference on Pattern Recognition*, pages 302–306.
- [Delon *et al.*, 2005] DELON, J., DESOLNEUX, A., LISANI, J.-L. et PETRO, A.-B. (2005). Color image segmentation using an acceptable histogram segmentation. *In Iberian Conference on Pattern Recognition and Image Analysis*, Istoril, Portugal.
- [Deng *et al.*, 1999] DENG, Y., MANJUNATH, B. et SHIN, H. (1999). Color image segmentation. *In International Conference on Computer Vision and Pattern Recognition*, Fort Collins, USA.
- [Desvignes, 1991] DESVIGNES, F. (1991). *Rayonnements optiques. Radiométrie-Photométrie*. Masson.
- [Dordet, 1990] DORDET, Y. (1990). *La colorimétrie. Principes et applications*. Eyrolles.
- [Drew *et al.*, 1998] DREW, M., WEI, J. et LI, Z.-N. (1998). On illumination invariance in color object recognition. *Pattern Recognition*, 31(8):1077–1087.
- [Drew *et al.*, 1999] DREW, M., WEI, J. et LI, Z.-N. (1999). Illumination-invariant image retrieval and video segmentation. *Pattern Recognition*, 32(8):1369–1388.
- [D’Zmura *et al.*, 1997] D’ZMURA, M., COLANTONI, P., KNOBLAUCH, K. et LAGET, B. (1997). Color transparency. *Perception*, 26:471–492.
- [D’Zmura *et al.*, 2000] D’ZMURA, M., RINNER, O. et GEGENFURTNER, K. (2000). The colors seen behind transparent filters. *Perception*, 29:911–926.
- [Espiau *et al.*, 2002] ESPIAU, F., MALIS, E. et RIVES, P. (2002). Robust features tracking for robotic applications : towards 2 1/2 D visual servoing with natural images. *In IEEE Int. Conf. on Robotics and Automation*, Washington, USA.
- [Faille, 2004] FAILLE, F. (2004). A fast method to improve the stability of interest point detection under illumination changes. *In IEEE International Conference on Image Processing*, pages 2673–2676, Singapour.
- [Fan *et al.*, 2001] FAN, J., YAU, D., ELMAGARMID, A. et AREF, W. (2001). Automatic image segmentation by integrating color edge extraction and seeded region-growing. *IEEE Transactions on Image Processing*, 10(10):1454–1466.
- [Faul et Ekroll, 2002] FAUL, F. et EKROLL, V. (2002). Psychophysical model of chromatic perceptual transparency based on subtractive color mixture. *Journal of the Optical Society of America A*, 19:1084–1095.

- [Ferri et Vidal, 1992] FERRI, F. et VIDAL, E. (1992). Colour image segmentation and labeling through multiedit condensing. *Pattern Recognition Letters*, 13(8):561–568.
- [Finlayson, 1995] FINLAYSON, G. (1995). *Coefficient Color Constancy*. Ph.d. thesis, Simon Fraser University.
- [Finlayson, 1996] FINLAYSON, G. (1996). Color in perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1034–1038.
- [Finlayson et al., 2000] FINLAYSON, G., HORDLEY, S. D., MARCHANT, J. et ONYANGO, C. (2000). Color invariance at a pixel. *In British Machine Vision Conference*, Bristol.
- [Finlayson et Schaefer, 2001] FINLAYSON, G. et SCHAEFER, G. (2001). Solving for colour constancy using a constrained dichromatic reflection model. *International Journal of Computer Vision*, 42(3):127–144.
- [Fontaine, 2001] FONTAINE, M. (2001). *Segmentation non supervisée d'images couleur par analyse de la connexité des pixels*. Thèse de doctorat. Université de Lille I.
- [Fontaine et al., 2000] FONTAINE, M., MACAIRE, L. et POSTAIRE, J.-G. (2000). Unsupervised segmentation based on connectivity analysis. *In International Conference on Pattern Recognition*, volume 1, pages 600–603, Barcelona, Spain.
- [Forgy, 1965] FORGY, E. (1965). Cluster analysis of multivariate data : efficiency versus interpretability of classification. *Biometrics*, 21:768–780.
- [Foucher et al., 2001] FOUCHER, P., REVOLLON, P. et VIGOUROUX, B. (2001). Segmentation d'images en couleur par réseaux de neurones : application au domaine végétal. *In Congrès ORASIS'01*.
- [Geusebroek et al., 2000a] GEUSEBROEK, J., KOELMA, D., SMEULDERS, A. et GEVERS, T. (2000a). Image retrieval and segmentation based on color invariants. *In IEEE International Conference on Pattern Recognition*.
- [Geusebroek et al., 2003] GEUSEBROEK, J.-M., BOOMGAARD, R. V. D., SMEULDERS, A. et GEVERS, T. (2003). Color constancy from physical principles. *Pattern recognition Letters*, 24:1653–1662.
- [Geusebroek et al., 2001] GEUSEBROEK, J.-M., BOOMGAARD, R. V. D., SMEULDERS, A. W. M. et GEERTS, A. W. M. (2001). Color invariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1338–1350.
- [Geusebroek et al., 2000b] GEUSEBROEK, J. M., SMEULDERS, A. W. M. et BOOMGAARD, R. V. D. (2000b). Measurement of color invariants. *In IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 50–57.
- [Gevers, 2001] GEVERS, T. (2001). Robust histogram construction from color invariants. *In IEEE International Conference on Computer Vision*, pages 615–620, Vancouver.
- [Gevers, 2002] GEVERS, T. (2002). Adaptive image segmentation by combining photometric invariant region and edge information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6):848–852.
- [Gevers, 2003a] GEVERS, T. (2003a). Reflectance-based classification of color edges. *In IEEE International Conference on Computer Vision*, volume 2, pages 856–861, Nice, France.

- [Gevers, 2003b] GEVERS, T. (2003b). Robust segmentation and tracking of colored objects in video. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(6):776–781.
- [Gevers et Groen, 1991] GEVERS, T. et GROEN, F. (1991). Segmentation of color images. *In Scandinavian Conference on Image Analysis*.
- [Gevers et Kajcovski, 1994] GEVERS, T. et KAJCOVSKI, V. (1994). Image segmentation by directed region subdivision. *In International Conference on Pattern Recognition*, volume A, pages 342–346.
- [Gevers et Smeulders, 1996] GEVERS, T. et SMEULDERS, A. (1996). A comparative study of several color models for color image invariant retrieval. *In Proc. 1st Int. Workshop on Image Databases and Multimedia Search*, pages 17–26, Amsterdam, Netherlands.
- [Gevers et Smeulders, 1997a] GEVERS, T. et SMEULDERS, A. (1997a). Object recognition based on photometric color invariants. *In Scandinavian Conference on Image Analysis*, Lappeenranta, Finland.
- [Gevers et Smeulders, 1997b] GEVERS, T. et SMEULDERS, A. W. M. (1997b). Edge steered region segmentation by photometric color invariant. *In Scandinavian Conference on Image Analysis*, Lappeenranta, Finland.
- [Gevers et Smeulders, 2000] GEVERS, T. et SMEULDERS, A. W. M. (2000). Pictoseek : Combining color and shape invariant features for image retrieval. *IEEE Transactions on Image Processing*, 9(1):102–119.
- [Gevers et al., 1998] GEVERS, T., SMEULDERS, A. W. M. et STOKMAN, H. M. G. (1998). Photometric invariant region detection. *In The Ninth British Machine Vision Conference*, Southampton UK.
- [Gevers et Stokman, 2003] GEVERS, T. et STOKMAN, H. M. G. (2003). Classifying of color edges in video into shadow-geometry, highlight, or material transitions. *IEEE Transactions on Multimedia*, 5(2):237–243.
- [Golland et Bruckstein, 1997] GOLLAND, P. et BRUCKSTEIN, A. M. (1997). Motion from color. *Computer Vision and Image Understanding*, 68(3):346–362.
- [Gouet et al., 2000] GOUET, V., MONTESINOS, P., DERICHE, R. et PELÉ, D. (2000). Evaluation de détecteurs de points d'intérêt pour la couleur. *In Reconnaissance de Formes et Intelligence Artificielle*, pages 257–266, Paris.
- [Gouiffès et al., 2004a] GOUIFFÈS, M., COLLEWET, C. et and A. TRÉMEAU, C. F.-M. (2004a). Color segmentation of inked characters. application to meat traceability control. *In IEEE International Conference on Image Processing*, pages 195–198, Singapour.
- [Gouiffès et al., 2004b] GOUIFFÈS, M., COLLEWET, C. et and A. TRÉMEAU, C. F.-M. (2004b). Invariants couleur dédiés à la transparence de l'encre. application à la segmentation de marquages. *Traitement du Signal*, 21(5):439–452.
- [Gouiffès et al., 2005] GOUIFFÈS, M., COLLEWET, C., FERNANDEZ-MALOIGNE, C. et TRÉMEAU, A. (2005). Suivi de points robuste aux spéularités et aux changements d'éclairage. *In 20^e colloque GRETSI sur le Traitement du Signal et des Images*, Belgique.
- [Gouiffès et al., 2004c] GOUIFFÈS, M., MARTY-MAHÉ, P. et and A. TRÉMEAU, C. F.-M. (2004c). Comparaison de deux méthodes de segmentation couleur appliquées à la traçabilité de produits carnés. *In 14^e Congrès Francophone AFRIF-RFIA de Reconnaissance des Formes et Intelligence Artificielle*, Toulouse, France.

- [Gros, 2000] GROS, P. (2000). Color illumination models for image matching and indexing. *In International Conference on Pattern Recognition*, volume 3, pages 580–583, Barcelona, Spain.
- [Guo *et al.*, 1998] GUO, G., S.YU *et* MA, S. (1998). Unsupervised segmentation of color images. *IEEE International Conference on Image Processing*, pages 299–302.
- [Hager *et* Belhumeur, 1998] HAGER, G. D. *et* BELHUMEUR, P. N. (1998). Efficient region tracking with parametric models of geometry and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(10):1025–1039.
- [Harris *et* Stephens, 1988] HARRIS, C. *et* STEPHENS, M. (1988). A combined corner and edge detector. *In 4th Alvey Vision Conference*, pages 147–151.
- [Hartigan, 1975] HARTIGAN, J. (1975). *Clustering algorithms*. John Wiley and Sons.
- [Healey, 1989] HEALEY, G. (1989). Using color for geometry insensitive segmentation. *Journal of the Optical Society of America A.*, 6:920–937.
- [Healey, 1992] HEALEY, G. (1992). Segmenting images using normalized color. *IEEE Transactions on systems, man and cybernetics*, 22(1):64–73.
- [Healey *et* Binford, 1988] HEALEY, G. *et* BINFORD, T. (1988). Local shape from specularly. *Computer Vision Graphics Image Processing*, 42:62–86.
- [Heigl *et al.*, 1999] HEIGL, B., PAULUS, D. *et* NIEMANN, H. (1999). Tracking points in sequences of color images. *In German-Russian workshop Pattern Recognition and Image Understanding*, pages 70–77, Herrsching, Germany.
- [Helmholtz, 1866] HELMHOLTZ, H. (1866). *Handbuch der physiologischen optik*. Hambourg and Leipzig.
- [Holla, 1982] HOLLA, K. (1982). Opponent colors as a 2-dimensional feature within a model of the first stages of the human visual system. *In International Conference on Pattern Recognition*, pages 561–563, Munich, Germany.
- [Horaud *et al.*, 1990] HORAUD, R., SKORDAS, T. *et* VEILLON, F. (1990). Finding geometric and relational structures in images. *In European Conference on Computer Vision*, pages 374–384.
- [Horn, 1986] HORN, B. (1986). Robot vision. *The MIT Press*.
- [Horn *et* Schunck, 1981] HORN, K. *et* SCHUNCK, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 7:185–203.
- [Horowitz *et* Pavlidis, 1974] HOROWITZ, S. *et* PAVLIDIS, T. (1974). Picture segmentation by a directed split-and-merge procedure. *In 2nd International Joint Conference on Pattern Recognition*, pages 424–433.
- [Huang, 1992] HUANG, C. (1992). Parallel image segmentation using modified HOPFIELD model. *Pattern Recognition Letters*, 13(5):345–353.
- [Huang *et al.*, 1992] HUANG, C., CHENG, T. *et* CHEN, C. (1992). Color images segmentation using scale space filter and markov random field. *Pattern Recognition*, 25(10):1217–1229.
- [Huber, 1981] HUBER, P.-J. (1981). *Robust statistics*. John Woley and Sons Inc., New York.
- [Hunter *et* Harold, 1987] HUNTER, R. *et* HAROLD, R. (1987). *The measurement of Appearance*. Wiley-Interscience.

- [Husenet, 2002] HUSSENET, L. (2002). *De l'analyse d'images à la reconstruction 3D : Application au contrôle d'objets métalliques*. Thèse de doctorat, Université de Reims, Champagne-Ardenne.
- [Ikonomakis *et al.*, 2000] IKONOMAKIS, N., PLATANIOTIS, K. et VENETSANOPOULOS, A. (2000). Color image segmentation for multimedia applications. *Journal Of Intelligent and Robotics Systems*, 28(1-2):5–20.
- [Imelinska *et al.*, 2000] IMELINSKA, C., DOWNES, M. et YUAN, W. (2000). Semi-automated color segmentation of anatomical tissue. *Computerized Medical Imaging and Graphics*, 3(24):173–180.
- [Jin *et al.*, 2001] JIN, H., FAVARO, P. et SOATTO, S. (2001). Real-time feature tracking and outlier rejection with changes in illumination. In *IEEE International Conference on Computer Vision*, pages 684–689, Vancouver, Canada.
- [Joly, 1999] JOLY, B. (1999). Stage de formation continue de l'École Technique Supérieure du Laboratoire.
- [Jones et Rehg, 1999] JONES, M. et REHG, J. (1999). Statistical models with application to skin detection. In *Proceedings IEEE Computer Vision and Pattern Recognition*.
- [Jurie et Dhome, 2002] JURIE, F. et DHOME, M. (2002). Hyperplane approximation for template matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):996–1000.
- [Kanade et Okutomi, 1994] KANADE, T. et OKUTOMI, M. (1994). A stereo matching algorithm with an adaptive window : theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):920–932.
- [Kass *et al.*, 1988] KASS, M., WITKIN, A. et TERZOPOULOS, D. (1988). Snakes : active contour models. *International Journal of Computer Vision*, 1:321–331.
- [Kazanov, 2004] KAZANOV, M. (2004). A new color image segmentation algorithm based on watershed transformation. In *International Conference on Pattern Recognition*, volume 2, pages 590–593.
- [Kermad *et al.*, 2002] KERMADE, C., COLLEWET, C. et CHEHDI, K. (2002). Apport de la sélection de primitives point à un suivi robuste. In *13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, volume 2, pages 539–548, Angers, France.
- [Kim *et al.*, 2004] KIM, D., LEE, K. et LEE, D. (2004). A novel initialization scheme for the fuzzy c-means algorithm for color clustering. *Pattern Recognition Letters*, 25(2):227–237.
- [Kim *et al.*, 2001] KIM, J., KIM, C., SEO, K. et KWEON, I. (2001). Color indexing using chromatic invariant. *Pattern Recognition*, 34(6):1189–1197.
- [Kim *et al.*, 2005] KIM, Y.-H., MARTÍNEZ, A. et KAK, A. (2005). Robust motion estimation under varying illumination. *Image and Vision Computing*, 23(4):365–375.
- [Kitchen et Rosenfeld, 1982] KITCHEN, L. et ROSENFELD, A. (1982). Gray-level corner detection. *Pattern recognition letters*, 1:95–102.
- [Klinker *et al.*, 1988] KLINKER, G., SHAFER, S. et KANADE, T. (1988). Image segmentation and reflection analysis through color. *SPIE. Applications of Artificial Intelligence VI*, 937:229–244.

- [Klinker *et al.*, 1990] KLINKER, G., SHAFER, S. et KANADE, T. (1990). A physical approach to image understanding. *International Journal of Computer Vision*, 4:7–38.
- [Koonty *et al.*, 1976] KOONTY, W., NARENDRA, P. et FUKUNAYA, F. (1976). A graph theoretic approach to non-parametric cluster analysis. *IEEE Transactions on Computer*, 25(9):961–973.
- [Kravtchenko et Little, 1999] KRAVTCHENKO, V. et LITTLE, J. (1999). Efficient color object segmentation using the dichromatic model. In *IEEE Pacific RIM Conference on Communication, Computer and Signal Processing*, pages 90–94.
- [Kurugollu *et al.*, 2001] KURUGOLLU, F., SANKUR, B. et HARMANCI, A. (2001). Color image segmentation using histogram multithresholding and fusion. *Image and Vision Computing*, 19(13):915–928.
- [Laganier, 1998] LAGANIERE, R. (1998). A morphological operator for corner detection. *Pattern Recognition*, 31(11):1643–1652.
- [Lambert, 1760] LAMBERT, J. (1760). *Photometria sive de mensura de gradibus luminis, colorum et ombrage*. Ebrhard Klett, Augsburg.
- [Lee et Cok, 1991] LEE, H. et COK, D. (1991). Detecting boundaries in vector field. *IEEE Transactions on Signal Processing*, 39:1181–1194.
- [Legéard, 1999] LEGÉARD, D. (1999). *Extraction de caractéristiques pour l'évaluation automatique de la qualité de jambons frais par vision numérique*. Thèse de doctorat, Thèse de doctorat. Université de Rennes I, Cemagref.
- [Lehman et Palm, 2001] LEHMAN, T. et PALM, C. (2001). Color line search for illuminant estimation in real-world scenes. *Journal of the Optical Society of America A*, 18(11):2679–2691.
- [Leydier *et al.*, 2004] LEYDIER, Y., BOURGEOIS, F. L. et EMPTOZ, H. (2004). Serialized k-means for adaptive color image segmentation application to document images and others. In *Document Analysis Systems*, pages 252–263.
- [Lezoray et Cardot, 2002a] LEZORAY, O. et CARDOT, H. (2002a). Cooperation of color pixel classification schemes and color watershed : a study for microscopic images. *IEEE Transactions on Image Processing*, 11(7):783–789.
- [Lezoray et Cardot, 2002b] LEZORAY, O. et CARDOT, H. (2002b). Segmentation of color images by clustering 2d histogram and merging regions. *Moroccan Journal of Control Computer Science and Signal Processing*.
- [Lezoray et Elmoataz, 2003] LEZORAY, O. et ELMOATAZ, A. (2003). Graph based smoothing and segmentation of color images. In *International Symposium on Signal Processing and its Applications*.
- [Lim et Lee, 1990] LIM, Y. et LEE, S. (1990). On the color image segmentation algorithm based on the thresholding and the fuzzy c-mean techniques. *Pattern Recognition*, 23(9):935–952.
- [Lin et Chen, 1991] LIN, X. et CHEN, S. (1991). Color image segmentation using modified HSI system for road following. *IEEE International Conference on Robotics and Automation*, pages 1998–2003.

- [Lin *et al.*, 2002] LIN, Z., WANG, J. et MA, K. (2002). Using eigencolor normalization for illumination-invariant color object recognition. *Pattern Recognition*, 35(11):2629–2642.
- [Lingrand, 1999] LINGRAND, D. (1999). *Analyse adaptative du mouvement dans les séquences monoculaires non calibrées*. Thèse de doctorat. Université de Sophia-Antipolis.
- [Lucas et Kanade, 1981] LUCAS, B. et KANADE, T. (1981). An iterative image registration technique. *In International Joint Conference on Artificial Intelligence*, pages 674–679, Vancouver, British Columbia.
- [Lucchese et Mitra, 2001] LUCCHESE, L. et MITRA, S. (2001). Color segmentation through independant anisotropic diffusion of complex chromaticity and lightness. *In IEEE International Conference on Image Processing*, volume 1, pages 746–749, Greece.
- [MacAdam, 1942] MACADAM, D. (1942). Visual sensitivities to color differences in daylight. 32:247–274.
- [MacKenna *et al.*, 2000] MACKENNA, S. J., JABRI, S., DURIC, Z. et WECHSLER, H. (2000). Tracking interacting people. *In 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 348–353.
- [Maeda *et al.*, 2000] MAEDA, J., ISHIKAWA, C., NOVIANTO, S., TADEHARA, N. et SUZUKI, Y. (2000). Rough and accurate segmentation of natural color images using fuzzy region-growing algorithm. *In International Conference on Pattern Recognition*, volume 3, pages 642–645.
- [Malladi *et al.*, 1994] MALLADI, R., SETHIAN, J. et VEMURI, B. (1994). Evolutionary fronts for topology independant shape modeling and recovery. *In European Conference on Computer Vision*, pages 3–13.
- [Marty-Mahé *et al.*, 2003] MARTY-MAHÉ, P., LOISEL, P. et BROSSARD, D. (2003). Color image segmentation to detect defects on fresh hams. *In 6th International Conference on Quality Control by Artificial Vision*, Gatlinburg, USA.
- [Masson *et al.*, 2004] MASSON, L., DHOME, M. et JURIE, F. (2004). Robust real time tracking of 3D objects. *In International Conference on Pattern Recognition*, pages 252–255.
- [Matsuoka *et al.*, 2002] MATSUOKA, H., TAKEUCHI, T., KITAZAWA, H. et ONOZAWA, A. (2002). Representation of pseudo interreflection and transparency by considering characteristic of human vision. *Computer graphics forum*, 21(3):503–510.
- [Matthews *et al.*, 2004] MATTHEWS, I., ISHIKAWA, T. et BAKER, S. (2004). The template update problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6): 810–815.
- [Maxwell, 1996] MAXWELL, B. (1996). *Segmentation and interpretation using multiple physical hypotheses of image formation*. Ph.d. thesis, Carnegie Mellon University, Pittsburgh.
- [Maxwell et Shafer, 1994] MAXWELL, B. et SHAFER, S. (1994). A framework for segmentation using physical models of image formation. *In IEEE International Conference on Computer Vision and Pattern Recognition*.
- [Maxwell et Shafer, 1996] MAXWELL, B. et SHAFER, S. (1996). Physics-based segmentation : moving beyond color. *In IEEE Conference on Computer Vision and Pattern Recognition*, pages 742–749.

- [Maxwell, 1860] MAXWELL, J. (1860). On the theory of compound colours and the relations of the colours of the spectrum. *In Proceedings of the Royal Society of London*, volume 10, pages 404–484.
- [Meas-Yedid *et al.*, 2004] MEAS-YEDID, V., GLORY, E., MORELON, A., PINSET, C., STAMON, G. et OLIVO-MARIN, J.-C. (2004). Automatic color space selection for biological image segmentation. *In International Conference on Pattern Recognition*, volume 3, pages 514–517.
- [Medioni et Yasumoto, 1987] MEDIONI, G. et YASUMOTO, Y. (1987). Corner detection and curve representation using cubic b-splines. *Computer Vision, Graphics and Image Processing*, 39(1):267–278.
- [Metelli, 1970] METELLI, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics*, 13:59–66.
- [Meyer, 1992] MEYER, F. (1992). Color image segmentation. *In International Conference on Image Processing Applications*, pages 303–306.
- [Moghaddamzadeh et Bourbakis, 1997] MOGHADDAMZADEH, A. et BOURBAKIS, N. (1997). A fuzzy region growing approach for segmentation of color images. *Pattern Recognition*, 6(30):867–881.
- [Mokhtarian et Mackworth, 1986] MOKHTARIAN, F. et MACKWORTH, A. (1986). Scale-based description and recognition of planar curves and two dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):34–43.
- [Montesinos *et al.*, 2000] MONTESINOS, P., GOUET, V., DERICHE, R. et PELÉ, D. (2000). Matching color uncalibrated images using differential invariants. *Image and Vision Computing*, 18:659–671.
- [Moravec, 1979] MORAVEC, H. (1979). Visual mapping by a robot rover. *In Proceedings of the 6th International Joint Conference on Artificial Intelligence*, pages 598–600.
- [Mukawa, 1990] MUKAWA, N. (1990). Estimation of shape, reflection coefficients and illuminant direction from images sequences. *In IEEE International Conference on Computer Vision*, pages 507–512.
- [Narasimhan *et al.*, 2003] NARASIMHAN, S., RAMESH, V. et NAYAR, S. (2003). A class of photometric invariants : separating material from shape and illumination. *In International Conference on Computer Vision*, pages 1387–1394.
- [Navon *et al.*, 2005] NAVON, E., MILLER, O. et AVERBUCH, A. (2005). Color image segmentation based on adaptive local thresholds. *Image and Vision Computing*, 1(23):69–85.
- [Nayar *et al.*, 1997] NAYAR, S., FANG, X. et TERRANCE, B. (1997). Separation of reflection components using color and polarization. *International Journal of Computer Vision*, 21(3): 163–186.
- [Nayar *et al.*, 1991] NAYAR, S., IKEUCHI, K. et KANADE, T. (1991). Surface reflection : Physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):611–634.
- [Negahdaripour, 1998] NEGAHDARIPOUR, S. (1998). Revised definition of optical flow : integration of radiometric and geometric cues for dynamic scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(9):961–979.

- [Nguyen et Smeulders, 2002] NGUYEN, H. et SMEULDERS, A. (2002). Template tracking using color invariant pixel features. *In IEEE International Conference on Image Processing*, pages 569–573, Rochester.
- [Nikolaev et Nikolayev, 2003] NIKOLAEV, D. et NIKOLAYEV, P. (2003). Linear color segmentation and its implementation. *Computer Vision and Image Understanding*, 94:115–139.
- [Noordam et Broek, 2000] NOORDAM, J. et BROEK, W. V. D. (2000). Geometrically guided fuzzy c-means clustering for multivariate image segmentation. *In International Conference on Pattern Recognition*, volume 1, pages 462–465.
- [Odobez et Bouthemy, 1995] ODOBEZ, J. et BOUTHEMY, P. (1995). Robust multiresolution estimation of parametric motion models. *International Journal of Visual Communication and Image Representation*, 6(4):348–365.
- [Ohba et Ikeuchi, 2000] OHBA, K. et IKEUCHI, Y. S. K. (2000). Appearance-based visual learning and object recognition with illumination invariance. *Machine Vision and Applications*, 12:189–196.
- [Ohlander *et al.*, 1978] OHLANDER, R., PRICE, K. et REDDY, D. (1978). Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing*, 8:313–333.
- [Ohta *et al.*, 1980] OHTA, Y., KANADE, T. et SAKAI, T. (1980). Color information for region segmentation. *In Computer Graphics and Image Processing*, volume 13, pages 222–241.
- [Oliver *et al.*, 1997] OLIVER, N., PENTLAND, A. et BERARD, F. (1997). Lafter : lips and face real-time tracker. *In IEEE Computer Vision and Pattern Recognition*, pages 123–129.
- [Osher et Rudin, 1988] OSHER, S. et RUDIN, L. (1988). Fronts propagation with curvature dependent speed : algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49.
- [Otha *et al.*, 1980] OTHA, Y., KANADE, T. et SAKAI, T. (1980). Color information for region segmentation. *Computer Graphics and Image Processing*, 13:222–241.
- [Pal et Pal, 1993] PAL, N. et PAL, S. (1993). A review on image segmentation techniques. *Pattern Recognition*, 26(9):1277–1294.
- [Panjwani et Healey, 1995] PANJWANI, D. et HEALEY, G. (1995). Markov random field models for unsupervised segmentation of textured color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(10):939–954.
- [Park *et al.*, 1998] PARK, S., YUN, I. et LEE, S. (1998). Color image segmentation based on 3-d clustering : Morphological approach. *Pattern Recognition*, 31(8):1061–1076.
- [Phong, 1975] PHONG, B.-T. (1975). Illumination for computer generated images. *Communications of the ACM*, 18(6):311–317.
- [Phung *et al.*, 2005] PHUNG, S., BOUZERDOUM, A. et CHAI, D. (2005). Skin segmentation using color pixel classification : analysis and comparison. *IEEE Pattern Analysis and Machine Intelligence*, 27(1):148–154.
- [Pos, 1989] POS, O. D. (1989). *trasparenze*. Padua, Icone.
- [Postaire, 1987] POSTAIRE, J. (1987). *De l'image à la décision*.

- [Pressigout et Marchand, 2005] PRESSIGOUT, M. et MARCHAND, E. (2005). Real-time planar structure tracking for visual servoing : a contour and texture approach. *In IEEE/RSJ Intelligent Robots and Systems*, volume 2, pages 1701–1706, Edmontons, Canada.
- [Prewitt, 1970] PREWITT, J. (1970). *Object enhancement and extraction*, pages 75–149. Academic Press.
- [R. Alferez, 1999] R. ALFEREZ, Y. W. (1999). Geometric and illumination invariants for object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(6):505–536.
- [Raffy, 1993] RAFFY, G. (1993). *Détermination de défauts sur des carcasses de dinde par segmentation couleur*. Rapport de stage du Cemagref de Rennes.
- [Raffy, 1999] RAFFY, G. (1999). *Vision numérique couleur appliquée à l'évaluation de la qualité de carcasses de dinde*, Thèse de doctorat. Université Blaise Pascal, Clermont Ferrand, France. Cemagref de Rennes.
- [Rohr, 2002] ROHR, K. (2002). Recognizing corners by fitting parametric. *International Journal of Computer Vision*, 9(3):213–230.
- [Roué, 2004] ROUÉ, M. (2004). *Segmentation couleur et reconnaissance de caractères numériques*. Rapport de DEA de l'Université de Rennes I. Cemagref de Rennes.
- [Saarinen, 1994] SAARINEN, K. (1994). Color image segmentation by a watershed algorithm and region adjacency graph processing. *International Conference on Image Processing*, pages 1021–1025.
- [Saber et al., 1997] SABER, E., TEKALP, A. et BOZDAGI, G. (1997). Fusion of color and edge information for improved segmentation and edge linking. *Image and Vision Computing*, 10(15):769–780.
- [Sangwine et Horne, 1998] SANGWINE, S. et HORNE, R. (1998). *The colour processing handbook*. Chapman and Hall.
- [Sapiro, 1997] SAPIRO, G. (1997). Color snakes. *Computer Vision and Image understanding*, 68(2):247–253.
- [Schettini, 1993] SCHETTINI, R. (1993). A segmentation algorithm for color images. *Pattern Recognition Letters*, 14:499–506.
- [Schmid et Mohr, 1997] SCHMID, C. et MOHR, R. (1997). Local greyvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5): 530–535.
- [Schmid et al., 2000] SCHMID, C., MOHR, R. et BAUCKHAGE, C. (2000). Evaluation of interest point detector. *International Journal of Computer Vision*, 37(2):151–172.
- [Seelen et Bajcsy, 1997] SEELEN, U. C. V. et BAJCSY, R. (1997). *Adaptive correlation tracking of targets with changing scale*, pages 313–322. Morgan Kaufmann.
- [Shafer, 1985] SHAFER, S. (1985). Using color to separate reflection components. *Color Research and Applications*, 10(4):210–218.
- [Sharma et Trussell, 1942] SHARMA, G. et TRUSSELL, H. (1942). Digital color imaging. *IEEE Transactions on Image Processing*, 32:247–274.

- [Shashua, 1997] SHASHUA, A. (1997). On photometric issues in 3d visual recognition from a single 2d image. *International Journal of Computer Vision*, 21(1-2):2448–2459.
- [Shen et Castan, 1992] SHEN, J. et CASTAN, S. (1992). An optimal linear operator for step edge detection. In *Computer Vision Graphics and Image Processing*, volume 54, pages 112–133.
- [Shi et Tomasi, 1994] SHI, J. et TOMASI, C. (1994). Good features to track. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 593–600, Seattle, Washington, USA.
- [Shih et Cheng, 2005] SHIH, F. et CHENG, S. (2005). Automatic seeded region growing for color image segmentation. *Image and Vision Computing*, 23(10):877–886.
- [Shih et al., 2005] SHIH, F., CHUANG, C.-F. et GADDIPATI, V. (2005). A modified regulated morphological corner detector. *Pattern Recognition Letters*, 26:931–937.
- [Sigal et al., 2000] SIGAL, L., SCLAROFF, S. et ATHITSOS, V. (2000). Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. In *IEEE Computer Vision and Pattern Recognition*, volume 2, pages 152–159.
- [Skarbek et A.Koschan, 1994] SKARBEEK, W. et A.KOSCHAN (1994). Colour image segmentation. a survey. Rapport technique, Institute for Technical Informatics, Technical University of Berlin.
- [Soriano et al., 2000] SORIANO, M., HUOVINEN, S., MARTINKAUPPI, B. et LAAKSONEN, M. (2000). Skin detection in video under changing illumination conditions. In *International Conference on Pattern Recognition*, volume 1, pages 839–842.
- [Spors et Rabenstein, 2001] SPORS, S. et RABENSTEIN, R. (2001). A real-time face tracker for color video. In *IEEE International Conference on Acoustics Speech and Signal Processings (ICASSP)*, Utah, USA.
- [Stern et Efros, 2005] STERN, H. et EFROS, B. (2005). Adaptive color space switching for tracking under varying illumination. *Image and Vision Computing*, 23(3):353–364.
- [Tadjine et Joubert, 2003] TADJINE, H. H. et JOUBERT, G. (2003). Colour object tracking without shadow. In *The 23rd Picture Coding Symposium*, pages 391–394.
- [Takahashi et al., 1995] TAKAHASHI, K., NAKATANI, H. et ABE, K. (1995). Color image segmentation using ISODATA clustering method. *Asian Conference on Computer Vision*, 1:523–527.
- [Tan et Ikeuchi, 2005] TAN, R. et IKEUCHI, K. (2005). Separating reflection components of textured surfaces using a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):178–193.
- [Tissainayagam et Suter, 2004] TISSAINAYAGAM, P. et SUTER, D. (2004). Assessing the performance of corner detectors for point feature tracking applications. *Image and Vision Computing*, 22(6):663–679.
- [Tomasi et Kanade, 1991] TOMASI, C. et KANADE, T. (1991). Detection and tracking of point features. Technical report CMU-CS-91-132, Carnegie Mellon University.
- [Tominaga, 1986] TOMINAGA, S. (1986). Color image segmentation using three perceptual attributes. *IEEE Computer Vision and Pattern Recognition*, pages 628–630.

- [Tommasini *et al.*, 1998a] TOMMASINI, T., FUSIELLO, A., ROBERTO, V. et TRUCCO, E. (1998a). Robust feature tracking in underwater video sequences. *In Proceedings of IEEE Oceans'98*, pages 46–50, Nice, France.
- [Tommasini *et al.*, 1998b] TOMMASINI, T., FUSIELLO, A., TRUCCO, E. et ROBERTO, V. (1998b). Making good features to track better. *In IEEE Conference on Computer Vision and Pattern Recognition*, pages 178–183.
- [Tommasini *et al.*, 1999] TOMMASINI, T., FUSIELLO, A., TRUCCO, E. et ROBERTO, V. (1999). Improving feature tracking with robust statistics. *Pattern Analysis & Applications*, 2(4):312–320.
- [Torrance et Sparrow, 1967] TORRANCE, K. et SPARROW, E. (1967). Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9).
- [Trajkovic et Hedley, 1998] TRAJKOVIC, M. et HEDLEY, M. (1998). Fast corner detection. *Image and Vision Computing*, 16:75–87.
- [Trémeau et Borel, 1997] TRÉMEAU, A. et BOREL, N. (1997). A region growing and merging algorithm to color segmentation. *Pattern Recognition*, 30(7):1191–1203.
- [Trémeau *et al.*, 2004] TRÉMEAU, A., FERNANDEZ-MALOIGNE, C. et BONTON, P. (2004). *Image numérique couleur : de l'acquisition au traitement*. Dunod.
- [Trouvé, 1991] TROUVÉ, A. (1991). *La mesure de la couleur. Principes, techniques et produits du marché*. Cetim-Afnor.
- [Trémeau, 1998] TRÉMEAU, A. (1998). *Analyse d'Images Couleur : du pixel à la scène, Habilitation à Diriger des Recherches*. Université St-Etienne.
- [Vandenbroucke, 2000] VANDENBROUCKE, N. (2000). *Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football*. Thèse de doctorat. Université de Lille I, France.
- [Vandenbroucke *et al.*, 2003] VANDENBROUCKE, N., MACAIRE, L. et POSTAIRE, J. (2003). Color image segmentation by pixel classification in an adapted hybrid color space. application to soccer image analysis. *Computer Vision and Image understanding*, 90:190–216.
- [Vincent et Laganière, 2001] VINCENT, E. et LAGANIÈRE, R. (2001). Matching feature points in stereo pairs : a comparative study of some matching strategies. *Machine Graphics and Vision*, 10(3):237–259–932.
- [Vincent et Soille, 1991] VINCENT, L. et SOILLE, P. (1991). Watersheds in digital spaces : an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):499–506.
- [Wang et Singh, 2003] WANG, J. et SINGH, S. (2003). Video analysis of human dynamics - a survey. *Real-Time Imaging*, 9(5):320–345.
- [Wu et Huang, 2000] WU, Y. et HUANG, T. (2000). View-independant recognition of hand postures. *IEEE Conference on Computer Vision and Pattern Recognition*, 2:88–94.
- [Yang *et al.*, 2002] YANG, J., HAO, S. et CHUNG, P. (2002). Color image segmentation using fuzzy c-means and eigenspace projections. *Signal Processing*, 82(3):461–472.
- [Yang et Liu, 1994] YANG, Y. et LIU, J. (1994). Multiresolution image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(16):689–700.

- [Yeh, 2003] YEH, C.-H. (2003). Wavelet-based corner detection using eigenvectors of covariance matrices. *Pattern Recognition Letters*, 24:2797–2806.
- [Zenko, 1986] ZENZO, S. D. (1986). A note on the gradient of multi-image. *Computer Vision Graphics Image Process*, 33:116–125.
- [Zhang *et al.*, 1999] ZHANG, R., TSAI, P.-S., CRYER, J.-E. et SHAH, M. (1999). Shape from shading : a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8): 690–706.
- [Zhu *et al.*, 1996] ZHU, S., LEE, T. et YUILLE, A. (1996). Region competition : unifying snakes, region growing, Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900.
- [Zucker, 1976] ZUCKER, S. (1976). Region growing : childhood and adolescence. *Computer Graphics and Image Processing*, 5:382–399.
- [Zugaj et Lattuati, 1998] ZUGAJ, D. et LATTUATI, V. (1998). A new approach of color images segmentation based on fusing regions and edge segmentations outputs. *Pattern Recognition*, 31(2):105–113.

Contributions of color and reflection models to extract and track features

The understanding of physical mechanisms such as transparency, shading, lighting or specular variations, can make easier the processing and the analysis of images, either by modeling these phenomena or by computing attributes that are invariant with respect to them.

First of all, we study the transparency of a colored ink printed on a colored support. We propose some color attributes that are invariant with respect to the support color and with respect to the quantity of ink used. The latter are used to carry out the segmentation process of ink markings, which mixes a classification and a region-growing approach. The proposed technique is applied to pork traceability, where an identifier has to be detected on the pieces of meat. In this application, we compare several segmentation methods and we show the relevance of the use of invariant attributes.

Then, a study on reflection models enables us to improve the robustness of point tracking in luminance images. Indeed, most techniques assume that the objects are lambertian, which means that the dynamic property of the observed scene is reduced to the geometric aspects and neglects the photometric ones. A few approaches propose to compensate for an affine model of the illumination variations in a small area around the point to be tracked. However, as far as we know, their theoretical validity is not clearly justified. Therefore, through a study on standard reflection models, we deduce the assumptions on which they are based. Moreover, we propose a point tracking approach based on a local photometric model, in which the illumination variations are approximated by a continuous function around the points to be tracked. Several experimentations are carried out on different objects, which undergo illumination variations or not, and prove the good robustness of our technique in comparison with the existing methods. In the same way, the tracking of larger area of the scene is improved by a more comprehensive photometric model.

We also tackle the problem of robustness of the point tracking in color images with respect to illumination variations. Firstly, the method developed in intensity images is extended to the case of color ones. The use of color invariants ensures a global correction with respect to lighting changes. Furthermore, we improve the use of most relevant invariants by using a local photometric model. Experiments on real images sequences are proposed to compare the different approaches. They show the improvement of robustness of these techniques with respect to illumination changes.

Keywords : Color segmentation, point tracking, reflection models, illumination variations, specular highlights, photometric invariants, transparency.

Apports de la couleur et des modèles de réflexion pour l'extraction et le suivi de primitives

Comprendre les mécanismes de transparence, d'apparition d'ombres, de variation d'éclairage, de spécularités, ne peut que faciliter la mise en œuvre de tâches de traitement et d'analyse d'images, soit en modélisant ces phénomènes soit en calculant des attributs qui leur sont invariants. Cette thèse s'intéresse à deux problématiques liées à ces aspects.

Premièrement, nous étudions la transparence d'une encre colorée imprimée sur un support de couleur. Nous proposons des attributs couleur invariants vis-à-vis de la couleur du support et vis-à-vis de la quantité d'encre. Ceux-ci sont exploités dans la mise en œuvre d'un algorithme de segmentation de marquages à l'encre, alliant une classification et une croissance de région. Finalement, l'approche proposée est appliquée dans le cadre de la traçabilité de la viande porcine, où un identifiant doit être détecté sur des pièces de porc. Dans ce contexte applicatif, nous comparons plusieurs méthodes de segmentation mises en œuvre et montrons la pertinence de l'utilisation d'attributs invariants.

Ensuite, l'étude de modèles de réflexion a permis d'améliorer la robustesse du suivi de points d'intérêt. En effet, la plupart des techniques suppose que les objets sont lambertiens, admettant ainsi que la dynamique d'une scène se résume à l'aspect géométrique et non photométrique. Quelques approches proposent une compensation d'un modèle affine des variations d'illumination dans un voisinage proche du point à suivre, mais, à notre connaissance, leur fondement n'a pas été clairement justifié. Ainsi, à partir de l'étude des modèles de réflexion standards, nous déduisons les hypothèses sur lesquelles elles se basent. Ensuite, nous proposons une méthode de suivi de points basée sur un modèle photométrique local, les variations d'illumination étant approximées par une fonction continue autour du point à suivre. Des expérimentations sur divers types d'objets, soumis à des changements d'illumination ou non, prouvent la bonne robustesse de cette approche en comparaison aux méthodes existantes. De la même manière, le suivi de motifs plus larges est amélioré par un modèle photométrique plus complet.

Nous abordons également le problème de la robustesse du suivi de points couleur vis-à-vis des variations d'illumination. D'une part, la méthode développée dans les images d'intensité a été étendue au cas de la couleur. Ensuite, l'utilisation d'invariants couleur permet d'assurer une correction globale vis-à-vis des changements d'éclairage. Enfin, nous améliorons l'utilisation des invariants les plus pertinents par un modèle photométrique local. Des expériences sur séquences réelles permettent de comparer les différentes approches et de prouver la meilleure robustesse de ces techniques vis-à-vis des variations d'illumination.

Mots-clés : Segmentation couleur, suivi de points, modèles de réflexion, variations d'illumination et de spécularité, invariants photométriques, transparence.