UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE 1 U.F.R. d'Informatique, Électronique, Électrotechnique et Automatique Laboratoire d'Automatique LAGIS.

Numéro attribué par la bibliothèque : \_\_\_\_

# <u>THÈSE</u>

pour obtenir le grade de

# DOCTEUR DE L'UNIVERSITÉ DE LILLE 1

## Discipline : Automatique et Informatique Industrielle

Présentée et soutenue publiquement

par

Laurent BUSIN

le 8 décembre 2006

# Sélection automatique d'espaces couleur pour la segmentation d'images - Application au contrôle qualité de motifs de verres sérigraphiés.

## JURY :

JG. Postaire	Professeur Émérite à l'U.S.T.L .	Président
L. Brun	Professeur à l'E.N.S.I. de CAEN	Rapportei
F. Meriaudeau	Professeur à l'I.U.T. Le Creusot	Rapportei
A. Trémeau	Professeur à l'Université de Saint-Étienne	Examinat
L. Macaire	Maître de Conférences H.D.R. à l'U.S.T.L.	Co-Direct
N. Vandenbroucke	Enseignant-Chercheur à l'E.I.P.C.	Co-Direct
P. Tahon	Ingénieur à ARC International	Invité

Président de Jury Rapporteur Rapporteur Examinateur Co-Directeur de recherche Co-Directeur de recherche Invité

# Remerciements

J'ai intégré le laboratoire d'automatique LAGIS<sup>1</sup> en Juin 2003 pour y réaliser mon stage de DEA dans l'équipe Vision et Image dirigée par Monsieur **Jack-Gérard Postaire**. Les travaux effectués consistaient au contrôle d'aspect de verres ornés de motifs imprimés par sérigraphie. Ces derniers ont alors été poursuivis en thèse donnant lieu au travail présenté dans ce manuscrit.

Ces travaux de recherche ont dans un premier temps été encadrés par Messieurs Jack-Gérard Postaire et **Nicolas Vandenbroucke** puis par Monsieur **Ludovic Macaire**. Mon travail de thèse n'aurait pu avoir lieu sans l'aide de l'EIPC qui a intégralement financé ma bourse de thèse. Pour cette raison, je remercie tout particulièrement Monsieur **Emmanuel Leclercq** ancien directeur de l'EIPC<sup>2</sup> et Monsieur **Patrick Debay** responsable de la recherche à l'EIPC qui m'ont permis d'effectuer ces travaux de thèse et de débuter dans le métier d'enseignant-chercheur.

Mes travaux de recherche n'auraient été possibles sans le partenariat profond liant l'EIPC à la société **ARC International** puisque c'est pour les besoins de cette entreprise que s'est orienté l'aspect applicatif des travaux réalisés. Aussi, je remercie très vivement Monsieur **Pascal Tahon** d'abord pour avoir proposé l'application ainsi que pour m'avoir intégré dans le service vision du site de Arques qu'il dirige.

À ces remerciements doivent être associés :

- Monsieur Luc Brun et Monsieur Fabrice Mériaudeau pour avoir accepté d'être rapporteurs ainsi que pour leurs pertinentes remarques.
- Monsieur Alain Trémeau pour avoir consenti à être examinateur.
- L'ensemble de l'équipe Vision et Image pour la confiance qu'ils m'accordent puisque je suis en poste d'ATER à l'IUT de Lille en Génie Mécanique et Productique grâce à leurs soutiens, et plus particulièrement François Cabestaing pour m'avoir déchargé en enseignement durant ce premier semestre 2006 afin de pouvoir me consacrer entièrement

<sup>&</sup>lt;sup>1</sup>Laboratoire d'Automatique, Génie Informatique & Signal

<sup>&</sup>lt;sup>2</sup>École d'Ingénieurs du Pas-de-Calais

à l'aboutissement de mes travaux de thèse.

Je profite de cette page de remerciement pour souligner l'aide que **Jean Caron** et la société EURESYS m'ont apportés lors de l'élaboration du poste d'acquisition. Enfin, je tiens à témoigner ma gratitude aux personnes qui ont le plus permis de faire vivre cette aventure humaine qu'est la thèse :

- Nicolas Vandenbroucke car il est certainement celui qui a le plus contribué (après moi) à ce travail. Il fut un encadrant de thèse idéal et possède autant de qualités morales que d'endurance physique. Je préfère donc ne pas énumérer ces nombreuses qualités pour ne pas alourdir cette partie.
- Ludovic Macaire qui a succédé à Monsieur Postaire pour devenir mon directeur de thèse et qui c'est toujours énormément impliqué dans ce travail. Son aide et ses remarques furent toujours pertinentes et déterminantes. Il est d'une clairvoyance et d'une réactivité que je n'ai rencontrées que trop rarement.
- Si Monsieur Postaire est une de ces nombreuses personnes que l'on appelle Monsieur, il est surtout une de ces rares personnes pour laquelle le titre de Monsieur est la conséquence du respect qu'il inspire et non une simple civilité.
- Pascal Tahon pour sa jovialité, son entrain et son aide qu'il a toujours su m'accorder malgré son emploi du temps surchargé.
- À tous les collègues que j'ai pu côtoyer durant ces trois années de thèse, en particulier ceux qui m'ont directement aidé pour la réalisation de mes travaux comme Hervé
  Rossiny pour la programmation ou Frédéric Lagache pour les figures qu'il a réalisé. Je remercie aussi ceux avec qui j'ai travaillé au quotidien comme l'ensemble du personnel enseignant, administratif et technique de l'EIPC ainsi que les membres du service Vision de ARC International.
- Enfin, j'adresse une petite pensée à tous ceux qui me sont chers et que je ne citerai pas de peur d'en omettre.

# Table des matières

Remerc	iement	S		3
Introdu	ction g	énérale		9
Chapitr	e 1. A	cquisition	des images-ligne couleur	13
1.1	Acqui	sition d'in	ages par caméra linéaire couleur	15
	1.1.1	Principe	et définitions	15
	1.1.2	Définitio	on d'une image	17
	1.1.3	Calcul d	e la résolution transversale	18
	1.1.4	Calcul d	e la résolution longitudinale	18
	1.1.5	Bruits d'	un capteur CCD	19
1.2	Les sy	vstèmes d'é	éclairage	20
	1.2.1	Les diffé	rents types de réflexion	20
		1.2.1.1	La réflexion spéculaire	21
		1.2.1.2	La réflexion diffuse	21
		1.2.1.3	La réflexion mixte	22
	1.2.2	Les géor	nétries d'éclairage	22
		1.2.2.1	Éclairages direct et diffus	23
		1.2.2.2	Éclairage bright-field	23
		1.2.2.3	Éclairage co-axial	24
		1.2.2.4	Éclairage par dôme	25
	1.2.3	Les diffé	erents types de sources lumineuses	26
		1.2.3.1	Définition d'une source de lumière blanche	26
		1.2.3.2	Caractéristiques des sources de lumière	27
		1.2.3.3	Les sources fluorescentes	27
		1.2.3.4	Les diodes électroluminescentes	28
		1.2.3.5	Les sources à halogénure-métallique	29
		1.2.3.6	Les sources à halogène	29
	1.2.4	Synthèse	e sur les systèmes d'éclairage	29
1.3	Les ca	iméras liné	aires couleur	30
	1.3.1	Technolo	bgies utilisées pour l'obtention d'images-ligne	30
	1.3.2	Architec	ture	32
	1.3.3	Modes d	acquisition	33
	1.3.4	Contrain	tes technologiques	35
		1.3.4.1	Décalage spatial	35
		1.3.4.2	Sensibilité des photo-récepteurs	38
	1.3.5	Contrain	tes physiques	39
		1.3.5.1	Orientation des capteurs de la caméra	39
		1.3.5.2	Déplacement du produit à inspecter	41

		1.3.5.3 Orientation des capteurs par rapport à la direction transversale	44
		1.3.5.4 Orientation des capteurs par rapport à la direction longitudinale	45
	1.3.6	Contraintes optiques des caméras linéaires couleur	46
	1.3.7	Synthèse sur les caméras linéaires couleur	47
1.4	Les car	rtes d'acquisitions linéaires couleur	47
	1.4.1	Modes de fonctionnement des cartes d'acquisition	48
	1.4.2	Mode ADR (Advanced Downweb Resampling)	49
	1.4.3	Traitements embarqués sur les cartes d'acquisition	50
	1.4.4	Remplissage de la mémoire image	51
	1.4.5	Avenir des cartes d'acquisition	51
1.5	Synthè	se	52
Charit			==
	Pannás	presentation et segmentation des images numeriques comeur	55
2.1			55
	2.1.1		50
	2.1.2	Les espaces luminance-chrominance	59
		2.1.2.1 Les espaces perceptuellement uniformes	39
		2.1.2.2 Les espaces antagonistes	61
		2.1.2.3 Les espaces de television	63
	010	2.1.2.4 Les autres espaces	63
	2.1.3		64
		2.1.3.1 Les espaces de coordonnees cylindriques	64
	014	2.1.3.2 Les espaces de coordonnees perceptuelles	66
	2.1.4		69
		2.1.4.1 Analyse en composantes principales et espace d'Onta	69 70
	015	2.1.4.2 Analyse en composantes independantes	/0
	2.1.5		/0
2.2	2.1.6	Synthese sur les espace couleurs	/1
2.2	Segme	ntation des images numeriques couleur	13
	2.2.1	Segmentation d'images couleurs par analyse spatiale	15
	2.2.2	Segmentation par classification de pixels	77
		2.2.2.1 Approche métrique	77
		2.2.2.2 Approche statistique	78
		Approche marginale	79
• •	a	Approche vectorielle	80
2.3	Segme	ntation et espace couleur	82
2.4	Synthè	se	82
Chapiti	re 3. Se	gmentation par multi-seuillage	85
3.1	Calcul	d'histogrammes 1D couleur	86
3.2	Filtrag	e d'histogrammes 1D par lissage Gaussien	88
3.3	Détecti	ion des seuils	91
	3.3.1	Définition des seuils	91
		3.3.1.1 Détection des minima locaux	92
		3.3.1.2 Détection des points d'inflexion	93
	3.3.2	Influence du choix des seuils pour la segmentation d'une image en ni-	2
		veaux de gris	96
	3.3.3	Influence du choix des seuils pour la segmentation d'une image couleur	98
3.4	Constr	uction d'une classe	00

	3.5	Lissage	multi-éch	nelles d'histogrammes 1D		. 101
		3.5.1	Lissage C	Gaussien heuristique		. 101
			3.5.1.1	Principe		. 102
			3.5.1.2	Lissage global		. 103
			3.5.1.3	Lissage local		. 104
		3.5.2	Scale-spa	ce filtering		. 106
			3.5.2.1	Construction de l'empreinte		. 107
			3.5.2.2	Construction des nœuds		. 110
			3.5.2.3	Sélection des nœuds actifs		. 112
			3.5.2.4	Limites du critère de sélection des nœuds actifs		. 113
		3.5.3	Spatial Se	cale-space Filtering		. 116
			3.5.3.1	Le degré de connexité		. 116
			3.5.3.2	Le degré d'homogénéité		. 117
			3.5.3.3	Le degré de compacité		. 118
			3.5.3.4	Définition du critère de sélection des nœuds actifs		. 118
		3.5.4	Résultats	de l'image de synthèse		. 119
	3.6	Résulta	ts expérin	nentaux		. 121
	3.7	Conclu	sion			. 122
Cha	pitre	e 4. Sél	ection au	tomatique d'espaces couleur pour la segmentation d'in	nage	es 125
2	4.1	Méthoo	le de segn		•••	. 125
2	4.2	Déterm	ination de	s histogrammes $ID$	•••	. 127
2	4.3	Lissage	des histo	grammes	•••	. 129
4	4.4	Détection des seuils			. 130	
4	4.5	Sélectio	on de l'esp	bace couleur le plus discriminant	• •	. 131
		4.5.1	Pouvoir c	liscriminant basé sur la répartition spatiale des pixels	•••	. 132
		4.5.2	Sélection	de l'espace couleur		. 135
4	4.6	Constru	$d^{\prime}$		•••	. 137
4	4./	Extract	ion des pi	xels assignes a la classe construite	•••	. 139
4	4.8	Critere	d'arret.		•••	. 140
4	4.9	Resulta	ts		•••	. 141
		4.9.1	Image "n	ouse"	• •	. 142
		4.9.2	Image "p	eppers	•••	. 143
		4.9.3	Image je		• •	. 140
	4 10	4.9.4 Canala	Synthese		•••	. 148
2	4.10	Conclu	sion		•••	. 148
Cha	nitre	5. An	nlication	à la détection de défauts sur verres serigraphiés		151
	5_1	Présent	ation du p	oste d'acquisition		152
-		5.1.1	La camér	a linéaire couleur		. 153
		5.1.2	Le systèr	ne d'éclairage		. 154
		5.1.3	Le choix	de la focale		. 156
4	5.2	Présent	ation des	défauts		. 156
4	5.3	Contrôl	e d'aspec	t par analyse d'images couleur		. 159
		5.3.1	Procédur	e hors-ligne		. 161
			5.3.1.1	Acquisition d'une image de référence		. 163
			5.3.1.2	Sélection de la ROI de référence		. 163
			5.3.1.3	Saisie du nombre de classes		. 164
			5.3.1.4	Calcul des histogrammes 1D et des histogrammes lissés		. 164

	5.3.1.5	Sélection de l'espace couleur le plus discriminant	165
	5.3.1.6	Construction d'une classe de pixels	167
	5.3.1.7	Extraction des pixels assignés à la classe construite	167
	5.3.1.8	Critère d'arrêt	167
5.3.2	Procédu	re en-ligne	168
	5.3.2.1	Acquisition d'un verre à inspecter	169
	5.3.2.2	Recherche des ROI	169
	5.3.2.3	Cas mono-modal	170
	(	Calcul des vecteurs gradients des composantes $R, V$ et $B$	170
	F	Fusion des 3 vecteurs gradients	170
	5.3.2.4	Cas multi-modal	170
	(	Conversion de la ROI dans l'espace couleur le plus discriminant	170
	(	Construction et extraction d'une classe de pixels	171
	(	Comparaison des pixels assignés aux classes	171
	5.3.2.5	Décision	172
5.4 Résulta	ats expérin	mentaux	173
Conclusion gén	érale		177
Bibliographie			181
Table des figure	es		189
Table des tablea	aux		195

# Introduction générale

Mes travaux de recherche effectués dans le cadre de la thèse de doctorat en automatique et informatique industrielle portent principalement sur la détermination d'espaces de représentation de la couleur adaptés à la segmentation d'images numériques couleur.

Ils ont été entrepris depuis 2003 au sein de l'équipe Vision et Image du Laboratoire LAGIS de l'Université de Lille 1 et ont été menés avec l'appui de l'École d'Ingénieurs du Pas-de-Calais (EIPC) qui a financé intégralement ma bourse de doctorat. Cette thèse fait l'objet d'une collaboration industrielle avec le groupe industriel Arc International, leader mondial des arts de la table.

Le problème soumis par ce groupe est le développement d'un poste de contrôle qualité de l'aspect de motifs sérigraphiés sur la surface de verres. Un écran en inox équipe les machines à sérigraphier. Ces écrans sont des pochoirs sur lesquels on appose de la peinture émaillée qui se fige immédiatement sur le verre. Les articles sont ensuite recuits afin de fixer définitivement le décor. Tout dysfonctionnement de ce dispositif provoque l'apparition de défauts d'aspect sur les motifs produits.

Le développement de systèmes de contrôle-qualité par vision industrielle de l'aspect de produits plats en défilement continu a fait l'objet de nombreuses thèses dirigées par Jack-Gérard Postaire durant les années 90 [Taner, Mac93, Fir97, Car00]. Ces travaux décrivent le développement matériel et logiciel de postes d'inspection utilisant des caméras linéaires en niveaux de gris. Ils ont abouti à la maîtrise des aspects méthodologiques du développement de tels postes d'inspection. Par conséquent, plusieurs docteurs ayant travaillé sur cette problématique ont implanté sur de nombreux sites industriels des systèmes automatiques de contrôle d'aspect par vision linéaire en niveaux de gris. Cette problématique dont le savoir-faire de l'équipe de Jack-Gérard Postaire est reconnue par la communauté scientifique, n'a alors plus fait l'objet de recherches actives.

Dans le cadre de mon projet d'ingénieur au sein de l'EIPC encadré par des membres de l'équipe de Jack-Gérard Postaire, nous avons tout d'abord développé un poste d'inspection des verres sérigraphiés utilisant une caméra linéaire en niveaux de gris. Cette caméra observe une ligne projetée sur la surface de chacun des verres inspectés dont la forme est uniquement cylindrique. La solution retenue consiste à mettre en rotation le verre inspecté selon son axe principal d'inertie. Pendant la rotation, la caméra linéaire observe une ligne de la surface du verre parallèle à l'axe principal d'inertie du verre. Les images-ligne acquises successivement

durant la rotation du verre permettent de représenter le développé de la surface du verre inspecté. Comme pour la majorité des postes d'inspection par vision industrielle, un soin particulier a été porté à la définition du dispositif d'éclairage et au choix de la caméra utilisée afin de mettre en évidence les défauts au sein des images acquises. Plus particulièrement, la mise au point de la géométrie du dispositif d'éclairage de la surface du verre et celle de la disposition spatiale de la caméra ont été délicates à ajuster. L'image du développé de la surface du verre inspecté est alors partitionnée en régions par des procédures classiques de segmentation afin de détecter d'éventuels défauts d'aspect apparaissant sur les motifs inspectés.

De cette étude, nous en avons conclu que l'exploitation de la seule information de luminance présente dans les images en niveaux de gris ne suffit pas toujours à mettre en évidence les régions représentant les défauts apparaissant sur les motifs colorés.

C'est pourquoi nous avons proposé d'exploiter l'information couleur présente dans les images acquises par une caméra CCD linéaire couleur. Cette étude est le champ d'application de nos travaux de recherche menés pendant mon doctorat.

Les caméras linéaires couleur atteignant des performances satisfaisantes en termes de qualité d'images et de fréquence d'acquisition ne sont apparues que très récemment sur le marché de la vision industrielle. Par conséquent, la communauté des chercheurs en vision industrielle maîtrise encore mal la prise en main de telles caméras couleur qui acquièrent des images-ligne selon des procédés spécifiques. Pour cette raison, le premier chapitre de ce manuscrit présente les principes généraux de l'acquisition d'images-ligne couleur. Nous y détaillons les quatre éléments intervenant pour l'acquisition de telles images à savoir :

- le dispositif d'éclairage en termes de nature d'illuminant et de géométrie de la lumière incidente,
- les propriétés de réflexion de la surface de l'objet observé par la caméra,
- les technologies de capture de l'information couleur des caméras linéaire couleur,
- les procédés de reconstitution d'images matricielles par les cartes d'acquisition dédiées aux caméras linéaires couleurs.

Dans le cadre de ce travail de doctorat, nous nous sommes attachés à développer une méthodologie originale de segmentation d'images couleur qui tente d'utiliser au mieux les propriétés colorimétriques des pixels pour partitionner en régions les images acquises.

La couleur de chaque pixel provenant d'une image-ligne est généralement représentée à partir de trois composantes couleurs qui sont les composantes rouge (R), verte (V) et bleue (B). Dans notre démarche, nous prenons comme hypothèse que les composantes couleur utilisées pour représenter la couleur définissent des points dans un espace couleur. Les points associés aux pixels d'une même région dans l'image forment un nuage de points dans l'espace couleur. Ce nuage est séparé de ceux formés par les points associés aux pixels qui représentent les autres régions. Nous supposons alors qu'il est possible d'identifier chaque nuage de points et donc de reconstruire les classes de pixels correspondant chacune à un nuage de points par des méthodes classiques d'analyse de données multidimensionnelles.

Les premiers résultats montrent que, dans de nombreux cas, il est nécessaire d'améliorer la segmentation en régions de l'image dont la couleur est codée dans l'espace couleur (R, V, B). En effet, la couleur d'un pixel peut être représentée en utilisant d'autres espaces couleur, chacun ayant des propriétés différentes. C'est pourquoi nous présentons dans la première partie du chapitre 2 les espaces couleur les plus fréquemment utilisés dans le cadre de l'analyse d'images couleur, en mettant à chaque fois en lumière leurs origines et leurs propriétés.

Dans la seconde partie du chapitre 2, nous décrivons succinctement les méthodes de segmentation existantes et abordons l'influence du choix de l'espace de représentation sur la qualité des résultats obtenus par les méthodes de segmentation d'images couleur. De nombreux auteurs ont étudié les rapports entre les espaces couleur et l'analyse d'images couleur [TFMB04]. Beaucoup ont soulevé le problème du choix délicat de l'espace couleur pour la segmentation d'images. Des éléments de réponse ont été notamment apportés par Vandenbroucke grâce à une analyse bas-niveau de l'image qui étudie le lien entre les espaces couleur et les performances d'un algorithme de segmentation [Van00]. Cette analyse tente de déterminer les composantes couleur les mieux adaptées à des problèmes spécifiques de segmentation. Cette démarche a été développée dans le cadre de la classification de pixels par apprentissage supervisé.

Nous proposons de nous placer dans un contexte non supervisé pour comparer les différents espaces couleur. La problématique que nous abordons est la sélection d'espaces couleur les mieux adaptés pour construire les classes de pixels en présence dans l'image.

Notre contribution s'inscrivant dans un contexte de construction des classes de pixels par apprentissage non supervisé, nous nous sommes dirigés vers une segmentation d'images par multi-seuillage des composantes couleur d'un espace. Le produit cartésien des seuils détectés selon chacune des composantes d'un espace couleur permet de définir des parallélépipèdes rectangles dans l'espace tri-dimensionnel. L'ensemble de ces parallélépipèdes sont examinés pour déterminer ceux qui constituent les domaines de définition des classes de pixels dans l'espace couleur considéré. La détection des seuils s'effectue par l'analyse des histogrammes monodimensionnels des trois composantes couleur associées à l'espace couleur.

Dans le chapitre 3, nous décrivons cette méthode de multi-seuillage en insistant sur les deux points clés de cette approche : le lissage de chacun des histogrammes mono-dimensionnels afin de mettre en évidence ses pics significatifs et la détection à proprement parler des seuils délimitant ses pics. Notre démarche s'inspire de la méthode appelée "scale-space filtering" qui opère conjointement le lissage adaptatif de l'histogramme et la détection des seuils. Le principal inconvénient de cette approche est la prise en considération exclusive de la distribution des niveaux des pixels. Comme la phase de détection des seuils néglige l'arrangement spatial des pixels dans l'image, elle peut parfois ne pas mettre en évidence des pics correspondant à des régions effectives dans l'image. Nous proposons une amélioration de cette approche en y intégrant une mesure sur les propriétés de connexité des pixels. Nous terminons ce chapitre en présentant l'apport en termes de qualité de segmentation de la prise en compte conjointe de la distribution des niveaux et de leur arrangement spatial dans l'image.

Ohlander tente de réduire l'influence de la sélection des seuils en proposant une analyse itérative des histogrammes mono-dimensionnels afin de construire les classes par multi-seuillages successifs [OPR78]. Il a montré que cette approche permet de mettre en évidence des pics significatifs qui ne sont pas distinguables en prenant en considération l'ensemble des pixels de l'image.

Nous proposons de nous inspirer de cette approche afin de construire les classes de pixels de manière itérative. Le chapitre 4 dévoile l'originalité de notre approche qui consiste en la sélection non supervisée de l'espace couleur le plus discriminant pour construire chacune des classes de pixels en présence dans l'image. A chaque itération de l'algorithme, nous mesurons la qualité de discrimination des classes de pixels en présence quand leurs couleurs sont codées dans différents espaces couleurs. L'algorithme compare les mesures du pouvoir de discrimination associées à chaque espace couleur et sélectionne le plus discriminant. Il détermine alors un domaine de définition d'une classe ayant la forme d'un parallélépipède rectangle dans cet espace couleur. Les pixels dont les points couleur sont inclus dans ce domaine sont assignés à cette classe et extraits de l'image à segmenter pour ne plus être considérés par les itérations suivantes.

Un soin particulier a été apporté à la définition de la mesure du pouvoir de discrimination d'un espace couleur de telle sorte que la classe de pixels construite à chaque itération corresponde à une région significative dans l'image. Cette mesure prend simultanément en considération les distributions des points dans l'espace couleur considéré et l'arrangement spatial des pixels caractérisés par ces couleurs dans l'image.

Dans le 5<sup>ème</sup> et dernier chapitre, nous montrons comment notre approche s'applique particulièrement bien à la segmentation d'images couleur représentant la surface de verres sérigraphiés afin de détecter d'éventuels défauts d'aspect. Ces verres sont souvent produits en petites séries pour répondre au mieux aux exigences des clients de l'industriel. Les motifs sérigraphiés des verres produits sur une même ligne de production sont donc fréquemment modifiés. Cette forte cadence de modifications des séries nécessite la mise en place de procédures non supervisées de segmentation d'images numériques couleur permettant de réduire le temps entre les changements de séries. Nous détaillons dans ce dernier chapitre le poste d'inspection de contrôle qualité de l'aspect des motifs de verres sérigraphiés ainsi que les différents types de défauts qui peuvent être issus d'un tel procédé de sérigraphie.

# **Chapitre 1**

# Acquisition des images-ligne couleur

La capacité de percevoir son environnement permet à l'homme d'accomplir un grand nombre d'activités. Le fait de remplacer l'observateur humain par une machine ayant la possibilité de voir permet l'automatisation de nombreuses tâches fastidieuses ou dangereuses (détermination de caractéristiques géométriques d'une pièce, contrôle de présence-absence d'éléments, vidéo-surveillance...). Les progrès techniques, tant au niveau des caméras que des systèmes informa-tiques de traitement d'images numériques, ont permis un élargissement du champ d'application de la vision artificielle. Cependant l'utilisation de l'information fournie par la couleur tarde à se démocratiser pour le contrôle d'aspect de produits manufacturés pour plusieurs raisons :

- La plupart des contrôles effectués actuellement en industrie ne nécessitent pas l'exploitation de la couleur. Un contrôle qualité par image numérique couleur génère donc un surcoût financier injustifié par rapport à un contrôle à l'aide d'une caméra monochrome.
- La manipulation d'images couleur nécessite un niveau d'expertise plus important que celui que requiert la manipulation d'images en niveaux de gris, tant pour la mise au point du poste d'acquisition que des algorithmes mis en place pour analyser automatiquement les images couleur.
- Bien souvent, les outils de traitements d'images couleur fournis dans les librairies de fonctions se limitent à l'emploi d'algorithmes adaptés aux images en niveaux de gris et n'exploitent qu'une seule composante couleur parmi les trois disponibles.
- Les outils de traitements d'images couleur ont la réputation auprès des industriels de souffrir d'un problème de robustesse des résultats par rapport aux outils utilisés en niveaux de gris et d'utiliser un nombre important de paramètres (comme la conversion d'espaces couleur par exemple).

Nous venons d'évoquer les principales raisons qui expliquent que le contrôle qualité par images numériques couleur n'est utilisé que par un nombre restreint d'applications de vision industrielle. Cependant, afin de rester compétitif face à la concurrence internationale, les entreprises essaient non seulement d'automatiser au maximum leurs lignes de production, mais aussi d'obtenir une qualité du produit de plus en plus élevée. Ce niveau d'exigence les poussent à effectuer des contrôles qualité de plus en plus pointus sur l'aspect des produits manufacturés en automatisant également leur poste de contrôle.

Le contrôle qualité est un domaine qui a très largement participé à l'essor de la vision artificielle car le contrôle automatique à l'aide d'une caméra peut être exhaustif et objectif dans le temps, contrairement au contrôle effectué par une inspection visuelle humaine qui s'avère subjective [NJ95, LRS<sup>+</sup>01]. Par ailleurs, la surface des produits manufacturés peuvent représenter des motifs colorés, où des défauts d'aspect peuvent être provoqués par des écarts de couleur. Ces applications, demandant l'utilisation d'une caméra numérique couleur, sont par exemple :

- l'évaluation de la qualité de carcasses de dinde [Raf99],
- l'inspection de dalles de marbres [Leb02],
- le contrôle du marquage de jambon [Gou05].

Ces applications nécessitent d'acquérir des images de haute qualité ainsi que de développer des algorithmes performants. Pour l'inspection d'articles en milieu industriel par caméra couleur, la prise en compte de l'environnement extérieur est par conséquent primordiale pour l'obtention d'une image couleur de qualité. Le système d'acquisition doit satisfaire les contraintes industrielles parfois sévères telles que l'emplacement et l'encombrement dans la ligne de fabrication, la température ambiante dans le poste d'inspection, les vibrations, les perturbations électromagnétiques et les poussières provoquées par le processus industriel, la facilité de maintenance du dispositif d'inspection... L'obtention d'une image correctement acquise permet d'éviter des corrections numériques de l'image et d'augmenter la robustesse du traitement. Ceci aura également pour conséquence de diminuer le temps de traitement et donc d'augmenter la cadence maximale d'articles inspectés. Cette propriété est essentielle afin que le poste de contrôle par vision artificielle ne devienne pas un goulot d'étranglement au niveau du process.

Le niveau d'expertise pour le développement d'un poste de contrôle est très élevé lorsqu'il s'agit d'acquérir des images avec une caméra linéaire couleur. En effet, la technologie utilisée pour acquérir des images de la surface de produits de grandes dimensions ou de formes cylindriques demandent une mise en oeuvre très délicate.

Un *système de vision* comporte différents éléments qui doivent coopérer entre eux, comme la caméra, le PC ou bien le système d'éjection des pièces considérées comme rebuts. Nous nous restreindrons dans cette partie à l'analyse des éléments d'un système de vision qui agissent pour l'obtention d'une image numérique. L'ensemble de ces éléments s'appelle *chaîne d'acquisition* et la réalisation physique d'une chaîne d'acquisition est appelée *poste d'acquisition*. La figure 1.1 montre les principaux éléments de la chaîne d'acquisition qui seront le fil conducteur de cette partie.



Figure 1.1 : Chaîne d'acquisition d'un système de vision.

Ainsi, après avoir introduit certaines notions générales et définitions relatives aux caméras linéaires nous présenterons les différents types de réflexion de la lumière et diverses géométries d'éclairage qui peuvent être employées pour l'acquisition d'images-ligne à partir d'une caméra linéaire couleur ainsi que les sources de lumière les plus courantes avec l'emploi d'une caméra linéaire couleur. Puis, nous aborderons l'architecture d'une caméra linéaire couleur afin de bien comprendre les spécificités requises par ce type de technologie. Enfin, nous présenterons le rôle des cartes d'acquisition.

# 1.1 Acquisition d'images par caméra linéaire couleur

L'objectif de cette première partie est d'introduire certaines notions générales relatives aux caméras linéaires ainsi que de définir certains termes qui seront utilisés dans la suite de ce manuscrit.

# **1.1.1 Principe et définitions**

Le principe de base de l'acquisition par une caméra linéaire consiste en la conversion des rayons lumineux<sup>1</sup> réfléchis par la scène à observer en un signal vidéo à l'aide d'un capteur de forme linéaire parfois désigné sous le terme de *barrette* dans la littérature. Ce capteur linéaire est constitué d'*éléments photo-sensibles*<sup>2</sup> qui sont aussi appelés *photo-sites* ou *photo-récepteurs*. Leur disposition linéaire ne permet que l'observation d'une ligne de la scène restituée sous la forme d'une image-ligne contrairement aux caméras matricielles qui permettent l'acquisition d'images matricielles. La figure 1.2 montre un exemple de ce type de capteur et met en évidence sa forme linéaire. La technologie la plus répandue pour réaliser ce type de caméras est actuellement la technologie CCD (Charge Couple Device) dont les photo-récepteurs produisent une tension croissante en fonction de l'intensité lumineuse qu'ils reçoivent.

Les caméras linéaires sont notamment utilisées pour le contrôle qualité de produits plats en défilement continu, de produits de grandes dimensions ou de produits possédant un axe

<sup>&</sup>lt;sup>1</sup>Ces rayons lumineux sont aussi appelés stimuli.

<sup>&</sup>lt;sup>2</sup>Parfois appelé pixels par abus de langage.



**Figure 1.2 :** Image d'un capteur linéaire couleur Kodak *KLI* – 2113, disponible à l'adresse *http://www.kodak.com/US/en/dpq/site/SENSORS/name/KLI-2113\_product*.

de révolution. Ce type de caméra permet d'obtenir une représentation bidimensionnelle de la surface des produits à inspecter grâce à leurs défilements.

La seconde dimension de l'image peut-être obtenue soit en déplaçant le produit à inspecter devant la caméra, soit en déplaçant la caméra devant le produit comme c'est le cas pour les images issues des scanners ou des photocopieuses. On distingue alors deux directions perpendiculaires :

– la direction longitudinale,

- la direction transversale.

La direction longitudinale ou direction axiale correspond à la direction de défilement de l'article à inspecter tandis que la direction transversale correspond à la direction selon laquelle est positionné le capteur de la caméra linéaire. La figure 1.3 présente ces directions pour un objet possédant un axe de révolution. Cette figure se compose de deux images montrant l'image acquise d'un verre à partir d'une caméra matricielle (figure 1.3(b)) et d'une caméra linéaire (figure 1.3(b)). La caméra linéaire permet de faire l'acquisition de toute la surface du verre sans déformation géométrique due au volume du verre contrairement à celle acquise par une caméra matricielle. Les directions longitudinale et transversale sont reportées sur l'image 1.3(a) de la figure 1.3 pour un produit avec un axe de révolution. Dans la partie 1.3.5, nous étudierons plus précisément les règles régissant les directions longitudinale et transversale.



(a) Juxtaposition d'images-ligne acquise par une caméra linéaire couleur observant un verre.



(b) Image matricielle acquise par une caméra matricielle couleur observant un verre.

Figure 1.3 : Images acquises par des caméras linéaire et matricielle couleur.

# **1.1.2 Définition d'une image**

Une image matricielle numérique couleur *I* correspond à une juxtaposition des images-ligne acquises pendant le défilement.

L'axe horizontal de l'image *I* correspond à l'axe transversal défini suivant la direction du capteur linéaire couleur tandis que l'axe vertical de l'image *I* correspond à l'axe longitudinal défini par la direction du défilement de l'article à contrôler par rapport à la caméra comme indiqué sur l'image 1.3(a). Ainsi, pour obtenir une image matricielle à partir d'images-ligne, il suffit de mettre en mémoire les images-ligne acquises successivement les unes à la suite des autres.

Une image matricielle numérique couleur *I* est représentée par trois signaux échantillonnés bidimensionnels à support et à valeurs bornées que nous notons  $C^k(x,y)$  où x et y sont les coordonnées du point *P* dans le plan image de taille  $X \times Y$  et k une composante couleur de l'espace couleur (R,V,B) et  $[x,y] \in \mathbb{N}^2$ ,  $0 \le x \le X-1$ ;  $0 \le y \le Y-1$  et  $k = \{R,V,B\}$ . C(x,y) représente ainsi la couleur de ce pixel.

# **1.1.3** Calcul de la résolution transversale

Pour déterminer le nombre de photo-éléments  $R_t$  du capteur qui définit la *résolution transversale*, il faut connaître à la fois la largeur  $L_x$  de l'article à inspecter et la dimension  $\delta_x$  du plus petit défaut devant être détecté suivant la direction transversale. Il suffit alors de diviser la largeur de l'article par la taille du plus petit défaut à détecter [Pos98], comme le montre la relation 1.1 :

$$R_t = C_{\text{sec}} \times \frac{L_x}{\delta_x},\tag{1.1}$$

où  $C_{sec}$  est un coefficient de sécurité qui doit être égal au moins à deux, afin de respecter le théorème de Shannon sur le sous-échantillonnage. De manière générale, il est recommandé empiriquement que ce plus petit défaut soit projeté sur par au moins cinq pixels, ce qui signifie qu'il faut utiliser un coefficient de sécurité  $C_{sec}$  supérieur ou égal à 5. Actuellement, les caméras linéaires couleur disponibles sur le marché de la vision industrielle sont équipées de capteurs comprenant jusqu'à 4096 photo-éléments.

# **1.1.4** Calcul de la résolution longitudinale

La *résolution longitudinale* correspond à la fréquence minimale d'acquisition des imagesligne. Elle s'évalue dans la direction de défilement de l'article à analyser. Pour déterminer la résolution longitudinale, il faut connaître la dimension du plus petit défaut  $\delta_y$  qui doit être détecté dans cette direction ainsi que la vitesse de déplacement  $v_y$  de l'article à analyser. Pour déterminer la résolution longitudinale qui correspond à la fréquence d'acquisition, il suffit de diviser la dimension  $\delta_y$  de ce plus petit défaut par la vitesse de déplacement de l'article à analyser par rapport à la caméra. Pour les raisons énoncées lors du calcul de la résolution transversale, on applique le même coefficient de sécurité qui permet en plus d'éviter le phénomène d'anamorphose. L'équation 1.2 montre le calcul de la résolution longitudinale, à savoir le nombre d'acquisitions d'images-ligne à effectuer par seconde :

$$F_l = C_{\text{sec}} \frac{v_y}{\delta_y}.$$
(1.2)

Dans de nombreuses applications de vision linéaire, la pièce à analyser est portée devant la caméra par un convoyeur. Ce convoyeur ne fonctionne pas obligatoirement à vitesse fixe, ce qui a pour conséquence de modifier la résolution longitudinale. Pour résoudre ce problème, on utilise un codeur incrémental qui permet de connaître la vitesse instantanée du convoyeur et ainsi conserver la résolution longitudinale des acquisitions quelque soit la vitesse du tapis. La figure 1.4 montre notamment un codeur incrémental ainsi que la *ligne de visée* qui représente la ligne de la scène observée par la caméra linéaire couleur.



Figure 1.4 : Poste d'acquisition d'images-ligne.

# 1.1.5 Bruits d'un capteur CCD

Il faut remarquer qu'il est délicat de faire l'acquisition de deux images identiques d'une même ligne sous les mêmes conditions d'éclairage à cause du bruit généré par le dispositif d'acquisition. Le bruit présent dans une image peut provenir de plusieurs sources classées en deux catégories suivant qu'elles sont d'origines temporelles ou spatiales [Kod03]. Par exemple, un bruit d'origine spatial peut être présent dans l'image lorsque les sensibilités des photorécepteurs varient au sein d'un même capteur, tandis qu'un bruit temporel sera introduit à cause de variations thermiques qui auront pour conséquence de modifier le niveau de référence des condensateurs dans lesquels les charges emmagasinées par les photo-récepteurs sont converties en une tension.

# **1.2 Les systèmes d'éclairage**

Le choix du système d'éclairage utilisé pour une application de vision industrielle est primordial [BW02, CKHD<sup>+</sup>97]. Ainsi, un système d'éclairage judicieux est parfois l'élément clé pour l'inspection des articles de production [LM01]. De plus, un éclairage mal adapté au produit à inspecter ne pourra jamais être complètement compensé par le traitement des images acquises [Raf99]. En règle générale, un système d'éclairage est utilisé par un système de contrôle d'aspect pour mettre en évidence les zones représentant les défauts à détecter.

L'élaboration d'un système d'éclairage nécessite l'étude de trois points. Le premier point concerne les caractéristiques optiques du matériau à inspecter. Nous détaillerons donc dans un premier paragraphe les principaux modèles de réflexion de la lumière utilisés lors de l'élaboration d'un système d'éclairage.

Le deuxième point est la géométrie du système d'éclairage qui correspond à la disposition des sources pour illuminer le produit à inspecter par rapport à la caméra. Il existe de nombreuses géométries d'éclairage préconisées suivant l'inspection considérée. Par exemple, la géométrie du système d'éclairage sera différente suivant que l'on cherche à détecter des défauts de surface ou les contours d'une pièce. Généralement, les caméras linéaires couleur sont utilisées pour des applications où les caméras linéaires monochromes ont montré leurs limites, c'est-à-dire lorsque le défaut ne peut pas être mis en évidence dans une image acquise par une caméra monochrome. Pour cette raison, une caméra linéaire couleur sera utilisée pour différencier les couleurs avec des niveaux de luminance identiques ou c'est le caractère chromatique qui permet de mettre en évidence l'apparition du défaut. Par conséquent, nous ne détaillerons que les géométries d'éclairage permettant de distinguer les couleurs en produisant le minimum de zone d'ombres sur l'objet à analyser.

Que ce soit pour une application industrielle utilisant une caméra linéaire couleur ou une caméra monochrome, l'étude des deux premiers points reste identique. Le troisième point à analyser est le choix de la source lumineuse. Sur ce point, les applications utilisant la couleur sont beaucoup plus contraignantes que celles exploitant une caméra monochrome. Ainsi, dans le dernier paragraphe, nous détaillerons différents types de sources de lumière utilisées pour une application de contrôle qualité industrielle.

# **1.2.1** Les différents types de réflexion

Les rayons lumineux provenant d'une source peuvent être partiellement ou intégralement réfléchis par l'objet observé par la caméra. Les rayons lumineux qui sont intégralement réfléchis proviennent de la *réflexion spéculaire* tandis que les rayons partiellement réfléchis sont issus de

la *réflexion diffuse*. On parle de *réflexion mixte* lorsque ces deux types de réflexion sont combiné entre elles.

#### 1.2.1.1 La réflexion spéculaire

Lorsque les rayons lumineux incidents notés  $R_i$  provenant d'une source lumineuse sont intégralement réfléchis dans une direction unique, symétrique à celle de la lumière incidente par rapport à la normale  $\vec{k}$  de la surface à inspecter comme le montre la figure 1.5, on parle alors de réflexion spéculaire. Les rayons lumineux réfléchis notés  $R_r$  appartenant à la réflexion spéculaire conservent les caractéristiques de la source lumineuse, à savoir la quantité d'énergie émie par intervalles de longueurs d'onde.



Figure 1.5 : Réflexion spéculaire.

# 1.2.1.2 La réflexion diffuse

Les rayons réfléchis  $R_r$  peuvent également se disperser dans plusieurs directions à cause de la rugosité du matériau et de la pénétration de la lumière dans le matériau dans le cas de la réflexion diffuse. La réflexion est dite totalement diffuse pour une surface Lambertienne lorsque la lumière est diffusée de manière uniforme dans toutes les directions (cf. figure 1.6). Le modèle Lambertien est le modèle le plus utilisé pour décrire les phénomènes de réflexion d'un matériau (comme le verre) qui laisse pénétrer profondément l'onde incidente dans l'objet [Bru02].



Figure 1.6 : Réflexion parfaitement diffuse.

## 1.2.1.3 La réflexion mixte

Pour la plupart des matériaux, les réflexions spéculaires et diffuses sont combinées de telle sorte que la réflexion diffuse se développe de façon préférentielle selon les directions voisines de celle de la réflexion spéculaire. Il se crée alors une lobe de diffusion avec un maximum d'intensité autour de cette direction préférentielle comme le montre la figure 1.7.



Figure 1.7 : Réflexion mixte.

Il est important d'estimer la nature de la réflexion du matériau à inspecter afin de définir un système d'éclairage qui permet la détection des défauts du produit à inspecter.

# 1.2.2 Les géométries d'éclairage

Selon le type de réflexion que l'on souhaite exploiter afin de mettre en évidence certaines parties à inspecter, différentes géométries d'éclairage peuvent être utilisées. Ces géométries diffèrent les unes des autres en fonction de la position de l'éclairage par rapport au produit, sa répartition, son intensité ou son homogénéité... Nous présentons dans les paragraphes suivants les géométries principalement utilisées par des dispositifs d'inspection couplés avec des caméras linéaires couleur.

## 1.2.2.1 Éclairages direct et diffus

La lumière provenant d'une source peut-être projetée directement sur l'objet analysé. On parle alors d'*éclairage direct*. Cette technique peut engendrer une répartition hétérogène de l'intensité de la lumière sur la ligne de visée comme le montre la figure 1.8(a). Par ailleurs, on peut intercaler entre la source et le produit un diffuseur translucide une plaque en plexiglas par exemple qui disperse et atténue la lumière rendant l'éclairage plus uniforme. On parle alors d'*éclairage diffus* comme le montre la figure 1.8(b).



(b) Éclairage diffus.

Figure 1.8 : Éclairage diffus et éclairage direct.

# 1.2.2.2 Éclairage bright-field

Cette technique d'éclairage consiste à placer la source dans l'axe optique de la caméra. Quand l'objet est un miroir plan, la source lumineuse est complètement réfléchie dans la ligne de visée de la caméra. La figure 1.9 illustre cette technique d'éclairage où  $\theta$  est l'angle entre la direction incidente et la normale de la surface à inspecter. Pour les éclairages bright-field, cet angle est proche de 0°.



Figure 1.9 : Éclairage bright-field.

Par opposition, la technique d'éclairage dark-field consiste à ajuster un angle  $\theta$  à une valeur proche de 90°. Cependant nous ne détaillerons pas cette technique qui permet de mettre en avant des défauts de surface par des reflets tels que les rayures. En effet, une caméra linéaire monochrome est alors suffisante pour analyser les reflets ainsi obtenus.

#### 1.2.2.3 Éclairage co-axial

Cette technique utilise un miroir semi-réfléchissant permettant d'aligner l'axe optique de la caméra avec l'axe de l'éclairage. L'emploi du miroir semi-réfléchissant permet à la fois d'éclairer la scène à partir des rayons lumineux provenant de la source lumineuse et de laisser passer les rayons lumineux provenant de la scène à observer.

Son principe est illustré sur la figure 1.10. On peut voir sur cette figure qu'une partie des rayons lumineux  $R_1$  provenant de la source de lumière est réfléchie par le miroir tandis que l'autre partie est transmise par ce miroir. Ainsi, les rayons lumineux  $R_2$  sont réfléchis par le miroir tandis que les rayons  $R_3$  sont transmis. Les rayons lumineux  $R_2$  sont alors les rayons incidents arrivant sur le produit à inspecter. Conformément aux différents types de réflexion énoncés dans le paragraphe 1.2.1, certains d'entre eux (les rayons lumineux  $R_4$ ) vont être réfléchis par le produit à inspecter et se diviser à nouveau sur le miroir semi-réfléchissant. Les rayons  $R_4$  vont soit être réfléchis (les rayons lumineux  $R_5$ ) soit être transmis par le miroir semi-réfléchissants (les rayons lumineux  $R_6$ ). Les rayons lumineux  $R_6$  vont alors exciter les éléments photo-sensibles de la caméra.



Figure 1.10 : Éclairage co-axial.

# 1.2.2.4 Éclairage par dôme

Cette technique dite *éclairage par dôme* ou *CDI* (pour Cloudy Day Illumination) consiste en un éclairage diffus et a été créée pour les applications les plus difficiles comme l'inspection de surfaces spéculaires complexes telles que les roulements à billes par exemple. Cette technique utilise une sphère d'intégration qui peut-être combinée avec un éclairage co-axial. Cette sphère d'intégration permet d'éclairer la pièce à inspecter suivant toutes les directions évitant ainsi de produire des zones ombrées et des reflets. Cette technique est appelée "cloudy day" car, comme lors d'une journée nuageuse, elle ne produit pas d'ombres [IMA05a]. Son principe est illustré sur la figure 1.11.



Figure 1.11 : Éclairage par dôme.

# **1.2.3** Les différents types de sources lumineuses

De la même façon que le choix de la géométrie de l'éclairage est déterminant pour mettre en évidence les défauts de la pièce à inspecter, le choix du type de source lumineuse doit être effectué avec soin. Différents types de sources lumineuses existent et nous présenterons dans les paragraphes qui suivent celles qui sont principalement utilisées avec des caméras linéaires couleur.

# 1.2.3.1 Définition d'une source de lumière blanche

Pour remédier à la non uniformité des niveaux des composantes couleur caractérisant les pixels représentant un objet dont la couleur est un blanc de référence, on opère une balance des blancs et une correction d'éclairage. Ces procédés, que nous détaillerons plus tard, consistent à ajuster des gains et des décalages pour chaque composante couleur des pixels. Ces gains appliqués aux pixels amplifient le bruit d'une manière d'autant plus forte que l'écart entre les niveaux des pixels d'un blanc de référence est grand. Ainsi, si l'emploi d'une balance des blancs et d'une correction d'éclairage sont nécessaires pour pallier la non uniformité de l'éclairage, celles-ci ont pour inconvénient d'amplifier le bruit présent dans l'image.

Pour restituer les couleur des motifs observés dans les images-ligne acquises, il est essentiel d'utiliser une source de lumière qui excite les 3 capteurs *R*, *V* et *B* de la caméra linéaire couleur de manière identique pour un objet blanc de référence. Une telle source de lumière est appelée *source de lumière blanche*. Certaines sources de lumière blanche correspondant à des conditions d'observation courantes ont été normalisées par la CIE (Commission Internationale de l'Éclairage) sous le nom *d'illuminants*. Certaines caméras sont calibrées pour l'emploi d'un illuminant et permettent l'utilisation de matrices de passage quand les conditions d'éclairage utilisées sont celles d'un autre illuminant [ATM06].

#### 1.2.3.2 Caractéristiques des sources de lumière

Les illuminants sont caractérisés par leur représentation spectrale relative d'énergie, notée  $S(\lambda)$ . Il s'agit d'une normalisation à 1 ou 100 de la répartition spectrale d'énergie pour une longueur d'onde particulière, en général  $\lambda = 560$  nm pour la plupart des illuminants. La figure 1.12 présente la représentation spectrale relative d'énergie pour l'illuminant  $D_{65}$  qui correspond à la moyenne des lumières durant la journée.



Figure 1.12 : Répartition spectrale relative d'énergie de l'illuminant D65 normalisé par la CIE.

Une source lumineuse peut également être caractérisée par sa température de couleur. Celleci représente la température à laquelle il faudrait porter un corps noir de Planck pour obtenir une représentation spectrale d'énergie identique à celle de la source. Cette température de couleur s'exprime en Kelvins (K). Les lumières "chaudes" correspondent à des couleurs jaune-rouge et ont une température de couleur basse de moins de 3000 K. Les lumières "froides" correspondent à des couleurs bleu-violet et ont une température de couleur élevée comprise entre 5000 K et 10000 K. Mises à part les courbes de sensibilités spectrales et les températures de couleur, les sources lumineuses diffèrent de part leurs durées de vie, leurs puissances lumineuses, la mise en forme de la lumière émise et bien sûr leurs coûts financiers.

Les sources de lumière qui sont alimentées de manière continue sont appelées *sources continues*, contrairement aux sources directement alimentées par une tension sinusoïdale. Les fréquences d'acquisition élevées des caméras linéaires couleur requièrent l'utilisation d'une source continue. Dans le cadre de notre application, nous avons été amenés à tester des lampes fluorescentes, des diodes électro-luminescentes, des lampes à halogénure-métallique et des lampes à halogène.

#### **1.2.3.3** Les sources fluorescentes

Les sources fluorescentes communément appelées néons émettent des radiations dans le domaine de l'ultraviolet par le passage d'électrons lors d'une décharge dans un gaz. Ces radiations sont transformées dans le domaine du visible selon le principe de fluorescence avec une poudre électro-luminescente, qui recouvre l'intérieur du tube du verre. Les néons ne sont pas des sources continues puisque la lumière n'est émise que lors d'une décharge. Elles nécessitent par conséquent l'emploi d'un ballast haute fréquence qui permet d'émettre cette décharge à une fréquence élevée<sup>3</sup>. Les fréquences d'acquisition d'une caméra linéaire pouvant atteindre 10 kHz, il est nécessaire d'utiliser des ballasts haute fréquence fonctionnant à une fréquence de 20 à 40 kHz [Zan02]. Les sources fluorescentes bénéficient d'une durée de vie estimée à 5000 heures [Zan02] et les températures de couleur disponibles pour les différents illuminants varient entre 2700 K à 6500 K [OSR04]. Ce type de source lumineuse permet d'éclairer une ligne d'un objet sur une grande longueur avec une bonne homogénéité, à savoir une intensité identique sur toute la ligne de visée de la zone éclairée par la partie centrale du néon. Les zones de l'objet éclairées par les extrémités du néon recevant moins de lumière que celles éclairées par la partie centrale, il est conseillé d'utiliser un néon de plus grande longueur que celle de l'objet à analyser ainsi que de disposer de réflecteurs pour diriger la lumière sur la zone à inspecter.

# 1.2.3.4 Les diodes électroluminescentes

Les diodes électroluminescentes (aussi appelées DEL ou LED) sont des sources de lumière de plus en plus utilisées par les applications de vision industrielle. Leurs petites tailles permettent de construire des systèmes d'éclairage de formes très variées. Cependant, la puissance lumineuse des LED de petites tailles est faible. Il existe des LED de plus grandes tailles avec une puissance lumineuse plus importante. Mais l'emploi de ces LED n'assure pas une intensité de lumière homogène sur la ligne de visée de l'objet à inspecter. Il faut donc trouver un compromis entre puissance lumineuse nécessaire et répartition homogène de la lumière émise par la source sur la ligne de visée. Pour remédier à ce problème, il existe des sources de lumière à base de LED qui focalisent la lumière afin d'assurer une intensité de lumière homogène sur la ligne de visée de l'objet à inspecter tout en conservant une puissance lumineuse importante. Les LED bénéficient d'une durée de vie estimée à 50000 heures et sont ainsi le type de source le plus rentable pour les applications de vision industrielle.

Les LED blanches utilisées pour les applications de vision industrielle nécessitant une caméra couleur ont une température de couleur de 5500 *K*. La technologie des LED est actuellement en pleine expansion et remplace progressivement les autres technologies standard comme les lampes fluorescentes dans les applications de vision industrielle grâce à leur grande durée de vie, leur faible coût d'achat, leur faible consommation d'énergie, leur constance d'intensité dans le temps et leur facilité d'utilisation [Gro04].

Des travaux de recherche récents ont permis de synthétiser une source de lumière blanche avec une répartition très proche des illluminants *D*50, *D*55, *D*65 et *D*75 définis par la CIE à partir d'un système à plusieurs LED [LLL06]. Il est aussi important de souligner que les couleurs émises par les LED sont sujettes à une certaine versatilité comme on peut le remarquer

<sup>&</sup>lt;sup>3</sup>On parle de néons hautes fréquences par abus de langage.

dans [LLL06]. Ce phénomène s'explique par le fait que la plupart des fabricants d'éclairage de vision industrielle utilisent des LED sans tri préalable [IMA05a].

#### 1.2.3.5 Les sources à halogénure-métallique

La lumière des lampes à halogénure-métallique est produite par un arc électrique entretenu entre deux électrodes. Les courbes de sensibilité spectrale de ces lampes présentent des raies. Celles-ci peuvent être "lissées" par l'adjonction d'éléments tels que des gaz de remplissage, des vapeurs métalliques, des poudres éléctroluminescentes ou des gaz halogènes [Raf99]. La lampe que nous avons utilisée possède une température de couleur de 4250 *K* pour une durée de vie de 2500 heures [EUR03].

#### 1.2.3.6 Les sources à halogène

Les lampes à halogène ont un spectre très proche de celui du corps noir (source de Planck). L'emploi d'un gradateur permet à ce type de lampe de régler leur puissance lumineuse et d'en modifier ainsi leur température de couleur. Les lampes halogènes possèdent une température de couleur limitée à 3000 *K* pour une durée de vie de 2000 heures [Raf99]. Ce type de lumière émise tend nettement vers le rouge. Ceci représente un inconvénient majeur avec l'utilisation d'une caméra linéaire couleur, car les photo-récepteurs recouverts d'un filtre rouge absorbent beaucoup plus de photons que les autres photo-récepteurs. Il en résulte une image fortement bruitée et difficilement exploitable pour des traitements ultérieurs.

# **1.2.4** Synthèse sur les systèmes d'éclairage

Après avoir défini les directions longitudinale, transversale ainsi que la notion d'imageligne, nous avons présenté les systèmes d'éclairage qui permettent l'inspection d'un produit manufacturé par caméra linéaire couleur. Pour cela, les systèmes d'éclairage ont été abordés suivant trois points :

- 1. Les différents types de réflexion du matériau à inspecter.
- La géométrie d'éclairage qui concerne la disposition des sources lumineuses par rapport au produit à inspecter afin que les stimuli reçus par la caméra couleur dépendent du type de réflexion que l'on souhaite mettre en évidence.
- 3. Les différents types de source lumineuses dont l'objectif principal consiste à stimuler de manière équivalente les photo-récepteurs *R*, *V* et *B* de la caméra linéaire couleur pour un objet blanc de référence.

En outre, lors des essais effectués avec les différentes sources de lumière, nous avons pu remarquer que l'acquisition d'images-ligne par caméra linéaire couleur nécessitait une grande puissance lumineuse, par rapport aux autres types de caméras utilisées en vision industrielle. Ce constat s'explique à notre avis à partir des deux raisons suivantes :

- Comme la fréquence d'acquisition d'une caméra linéaire couleur atteint plusieurs kilo hertz, le temps d'intégration des capteurs couleur *R*, *V* et *B* se fait sur une période très courte.
- La présence de filtres colorés sur les capteurs implique que les photo-récepteurs doivent être exposés plus longtemps que les photo-récepteurs d'une caméra monochrome sensible sur un domaine spectral plus important.

# **1.3** Les caméras linéaires couleur

Le composant principal d'une caméra linéaire couleur est le capteur photo-sensible. Ce capteur est constitué d'un ensemble de photo-récepteurs qui convertissent les stimuli couleur en un signal électrique. Cette information électrique peut-être soit transmise directement à la carte si la caméra utilisée est de type numérique ; soit convertie en un signal analogique, si la caméra utilisée est analogique. Si la caméra utilisée est de type analogique, l'information fournie par la caméra est envoyée à la carte d'acquisition qui numérisera à nouveau le signal analogique reçu. La chute constante des prix du matériel et les progrès réalisés dans les transferts de données font que les caméras linéaires couleur sont depuis peu des produits exploitables par les entreprises. C'est pour cette raison que la plupart des caméras linéaires couleur sont numériques. En outre, nous ne présenterons pas les différents standards de transmissions des données de la caméra vers la carte d'acquisition. En effet le standard Camera Link [Ass04] développé par les principaux constructeurs de caméras et de cartes d'acquisition est aujourd'hui le standard le plus utilisé pour le transfert des données des caméras linéaires couleur vers la carte d'acquisition même si ce secteur est actuellement en pleine effervescence.

Nous développerons dans cette partie les différentes technologies utilisées par les caméras linéaires pour obtenir une image couleur. Une fois la caméra linéaire couleur choisie, nous détaillerons l'architecture et les différents modes d'acquisition disponibles sur une caméra linéaire couleur en vue d'obtenir une image-ligne. Cette image peut souffrir de différents défauts liés à l'éclairage, à la technologie employée par la caméra, au mouvement de la pièce analysée, à l'orientation de la caméra ainsi qu'à l'objectif utilisé. Ainsi, nous proposons de détailler les différents défauts d'images liés à l'utilisation d'une caméra linéaire couleur selon qu'ils sont issus de *contraintes technologiques* de la caméra, des *contraintes physiques* ou des *contraintes optiques*.

# 1.3.1 Technologies utilisées pour l'obtention d'images-ligne

Comme le souligne Gevers, les caméras couleur et l'oeil humain se basent sur le principe de la théorie trichrome des récepteurs visuels de Young-Helmholtz [Gev00]. Ce principe met en évidence que n'importe quelle couleur peut être reproduite par un mélange ou *synthèse additive* de rouge, de vert et de bleu appelées *couleurs primaires*. Les capteurs étant sensibles sur un domaine plus grand que le domaine du visible, il suffit de placer un filtre coloré devant le

photo-récepteur pour que celui-ci ne soit sensible que sur un intervalle de longueurs d'onde déterminées. Aussi, pour que l'image acquise soit proche de l'image que l'oeil perçoit, les filtres utilisés sont de couleur rouge, vert et bleu. L'acquisition d'images numériques couleur utilise donc trois capteurs de sensibilités différentes pour reconstituer la couleur d'un stimulus par synthèse additive. A partir de ce procédé physique, il existe quatre techniques principales qui permettent de faire l'acquisition d'images couleur [Mat00] :

- La première méthode consiste à prendre 3 images d'un même objet à l'aide d'une caméra monochrome. Chaque acquisition se fait en plaçant un filtre différent devant la caméra. Cette méthode a l'avantage d'être simple de mise en oeuvre, économique et permet de choisir les filtres. Cependant, elle ne fonctionne que si la scène est statique et qu'aucune cadence de production n'est imposée.
- 2. Le principe de la deuxième méthode est de lire le signal par couche (en fonction de la profondeur) dans les photo-récepteurs car la profondeur des électrons excités dans le capteur est proportionnelle à la longueur d'onde. En effet, les photons à haute énergie comme ceux du bleu sont absorbés à la surface du capteur photo-sensible, tandis que les photons à basse énergie pénètrent plus profondément dans la couche de silicium du capteur photo-sensible avant d'être absorbés. Cependant, il n'existe aucune caméra linéaire à notre connaissance utilisant ce principe disponible dans le commerce. Cette technologie nommée Foveon X3 existe toutefois pour des caméras matricielles [Hub05].
- 3. La troisième méthode consiste à utiliser un système de prismes optiques permettant de projeter une même image sur trois capteurs différents recouverts par des filtres colorés. Grâce à ce système de prismes, un même élément de la scène observé est projeté sur le même photo-élément des trois capteurs *R*, *V* et *B* comme indiqué sur la figure 1.13(a).
- 4. La dernière méthode est utilisée par la plupart des caméras grand public et industrielles. Cette méthode consiste à placer des filtres colorés sur chaque élément photo-sensible. Ainsi, pour une caméra linéaire, on disposera 3 capteurs différents (un rouge, un vert, un bleu) placés parallèlement et séparés les uns des autres d'une centaine de microns environ comme le montre la figure 1.13(b) avec des lignes de visée spécifiques à chacun des capteurs couleur. Nous traiterons par la suite des conséquences provoquées par l'emploi de lignes de visées différentes des capteurs *R*, *V* et *B*. A titre d'illustration, nous avons représenté les zones aveugles entre les lignes de visée de manière exagérée sur la figure 1.13(b). Cette technologie est la plus répandue car elle revient moins chère que la technologie utilisant un système de prismes. Par la suite, l'emploi du terme *caméra linéaire couleur* sera utilisé pour les caméra dont les capteurs couleurs sont montés en parallèle et le terme *caméra linéaire à prismes* sera utilisé pour désigner les caméras équipées de trois capteurs montés sur un système optique à base de prismes. Nous ne détaillerons que le principe des caméras linéaires couleur car cette technologie très répandue est aussi celle utilisée dans notre application industrielle.



Figure 1.13 : Lignes de visée d'une caméra tri-linéaire avec et sans système de prismes.

# 1.3.2 Architecture

La période durant laquelle les photo-récepteurs accumulent des charges électriques en fonction de l'intensité lumineuse reçue s'appelle le temps d'intégration. A la fin de ce temps d'intégration succède un laps de temps durant laquelle la surface en silicium du capteur transmet les charges accumulées dans le registre à décalage auquel elle est reliée. Le temps durant lequel les charges emmagasinées sont envoyées dans les registres à décalage pour être transformées en un signal électrique est appelé temps de transfert. Le temps de transfert commence lorsque la caméra reçoit le signal de commande souvent appelé signal de *readout*. Durant le temps de transfert, les photo-récepteurs ne sont plus sensibles aux rayons lumineux et n'emmagasinent donc plus de charges grâce à un obturateur (ou shutter) électronique. Les niveaux de tension des photo-récepteurs transmis aux registres à décalage peuvent être amplifiés par un VGC (Variable Gain Control) et sont convertis en valeurs numériques grâce à un dispositif ADC (Analogic to Digital Converter) [BAS03]. Il transforme la tension en une valeur codée sur 8 bits<sup>4</sup> que nous appellerons niveau d'une composante. Les niveaux des composantes proviennent d'un des capteurs couleur rouge R, vert V ou bleu B. Les niveaux des composantes R, V et B sont ensuite stockées dans des piles FIFO (First In First Out) différentes pour chaque capteur couleur. En fonction des paramètres de la caméra, un micro-contrôleur permet de sélectionner les valeurs appropriées des composantes R, V et B dans les piles FIFO suivant leur ordre d'arrivée dans les piles. Lorsque les informations des composantes R, V et B sont disponibles dans les piles FIFO, celles-ci sont envoyées à la carte d'acquisition. Cette information envoyée à la carte correspond aux niveaux de couleur R, V et B d'un point image aussi appelé pixel d'une image-ligne. La figure 1.14 illustre l'architecture fonctionnelle d'une caméra linéaire couleur.

<sup>&</sup>lt;sup>4</sup>On trouve aussi des caméras qui codent l'information jusqu'à 12 bits.



Figure 1.14 : Architecture fonctionnelle d'une caméra linéaire couleur.

Une caméra linéaire couleur n'étant qu'un maillon de la chaîne d'acquisition, elle propose plusieurs modes de fonctionnement en accord avec les autres éléments de la chaîne d'acquisition. En outre, de part la technologie utilisée, les caméras provoquent l'apparition de fausses couleurs dans l'image acquise. Nous détaillerons dans les paragraphes suivants, les différents modes de fonctionnement d'une caméra linéaire couleur ainsi que les défauts images liés à l'emploi de la caméra linéaire couleur.

# **1.3.3** Modes d'acquisition

Une tâche importante lors de l'acquisition d'images-ligne est le respect de la résolution de l'image dans la direction longitudinale. Afin de pouvoir respecter cette résolution longitudinale, la plupart des caméras linéaires couleur du marché proposent deux modes de fonctionnement [TAT04, Per, DAL05, BAS03] :

- le mode *free-run* parfois appelé mode maître, la caméra génère elle-même son signal de déclenchement d'acquisition appelé *signal de contrôle interne*. La caméra est équipée de registres de commande où l'utilisateur peut ajuster la fréquence d'acquisition des imagesligne. La caméra délivre alors les images-ligne à la fréquence ainsi précisée.
- le mode *ExSync* parfois appelé mode esclave qui permet de déclencher l'acquisition à partir d'un événement extérieur à la caméra. Cet événement proviendra généralement d'un codeur incrémental en contact direct avec l'objet dont on veut faire l'acquisition via la carte d'acquisition.

En mode "free run", la caméra utilise deux signaux appelés "timer 1" et "timer 2". Le timer 1 renseigne sur la durée de l'état bas du signal de contrôle interne de la caméra tandis que le paramètre timer 2 renseigne sur la durée de l'état haut. La période du signal de contrôle interne est égale à la somme des périodes des deux timers. En mode free run, la caméra peut travailler selon deux modes différents :

- Free run edge-controlled : Dans ce mode, les éléments photo-sensibles de la caméra sont exposés durant toute la période du signal de contrôle généré par la caméra. Le signal readout et le transfert de charges se déclenchent à chaque front montant du signal généré comme le montre la figure 1.15.



Figure 1.15 : Mode de déclenchement "Free run edge-controlled" pour une caméra linéaire couleur.

- Free run programmable : Dans ce mode, les éléments photo-sensibles de la caméra sont exposés pendant que le signal de contrôle généré par la caméra est à l'état bas. Le signal readout et le transfert de charges se déroulent à chaque front montant du signal généré comme le montre la figure 1.16.



Figure 1.16 : Mode de déclenchement "Free run programmable" pour une caméra linéaire couleur.

En mode ExSync, la caméra peut travailler suivant trois modes différents :

*ExSync edge-controlled* : Dans ce mode, les éléments photo-sensibles de la caméra sont exposés durant toute la période du signal ExSync. Le signal readout a lieu au front montant de chaque signal ExSync. La figure 1.17 illustre ce mode de fonctionnement.



Figure 1.17 : Mode de déclenchement "ExSync edge-controlled" pour une caméra linéaire couleur.

*ExSync level-controlled* : Le temps d'exposition est déterminé par l'intervalle situé entre le front descendant du signal ExSync courant et le front montant du prochain signal Ex-Sync. Le signal readout et le transfert ont lieu lors de chaque front montant du signal ExSync.



Figure 1.18 : Mode de déclenchement "ExSync level-controlled" pour une caméra linéaire couleur.

- ExSync level-programmable : A chaque front montant du signal ExSync commence l'intégration des stimuli par les éléments photo-sensibles. Ceux-ci intègrent pendant un temps prédéfini par le timer appelé timer 1. A la fin du timer, le signal readout est envoyé et les charges sont transférées comme le montre la figure 1.19.



Figure 1.19 : Mode de déclenchement "ExSync level-Programmable" pour une caméra linéaire couleur.

# **1.3.4** Contraintes technologiques

L'utilisation d'une caméra linéaire couleur impose deux principales contraintes technologiques liées au capteur que nous préciserons dans les paragraphes suivants.

## **1.3.4.1** Décalage spatial

Comme les trois capteurs linéaires sont espacés d'une centaine de microns, chaque barrette couleur n'observe pas la même surface du produit à analyser au même instant comme le montre la figure 1.20. Cette contrainte technologique entraîne l'apparition de fausses couleurs dans l'image acquise connue sous le nom de *phénomène de décalage spatial*.

Afin d'illustrer ce phénomène, supposons que l'on veuille acquérir l'image d'une ligne noire sur un fond blanc avec une caméra linéaire couleur. Les figures 1.21(a), 1.21(b) et 1.21(c)montrent les trois instants durant lesquels les acquisitions de la ligne noire sont effectuées par les capteurs *R*, *V* et *B*. Le tableau 1.21(d) renseigne sur l'état des capteurs pour ces trois instants ainsi que sur la couleur correspondante perçue par la caméra. Par exemple, à l'instant *t* qui



Figure 1.20 : Lignes de visée des capteurs d'une caméra linéaire couleur.

correspond à la figure 1.21(a), la ligne noire passe devant le capteur rouge. La couleur correspondante perçue par la caméra est le cyan car seuls les capteurs des barrettes vertes et bleues sont sensibilisés. Ainsi, les valeurs du tableau montrent que la ligne noire acquise à trois instants différents correspondra dans l'image non pas à une ligne noire mais à la succession d'une ligne cyan, puis d'une ligne magenta et enfin d'une ligne jaune. Ces lignes colorées cyan, magenta, jaune correspondent aux fausses couleurs énoncées précédemment.



(a) Première acquisition à l'instant t: La ligne noire est présente uniquement sur la ligne de visée du capteur R.



(c) Troisième acquisition à l'instant  $t + 2\Delta t$ : La ligne noire est présente uniquement sur la ligne de visée du capteur *B*.



(b) Deuxième acquisition à l'instant  $t + \Delta t$ : La ligne noire est présente uniquement sur la ligne de visée du capteur V.

Instant	t	$t + \Delta t$	$t + 2\Delta t$	
Capteur R	0	255	255	
Capteur V	255	0	255	
Capteur B	255	255	0	
Couleur				

(d) Valeur des composantes couleurs des pixels des images lignes acquises aux instants t,  $t + \Delta t$  et  $t + 2\Delta t$  et leurs couleurs correspondantes.

Figure 1.21 : Phénomène de décalage spatial.

Pour éviter le phénomène de décalage spatial, il suffit de recombiner les lignes acquises à des instants différents par les capteurs R, V et B empilés dans les piles *FIFO*. Deux paramètres de la caméra permettent de recombiner entre-elles des images-ligne acquises à des moments différents. Le premier paramètre indique quel est le premier capteur couleur rencontré par le défilement de l'objet à analyser et le second paramètre indique le décalage (en nombre de lignes) souhaité pour recaler les acquisitions des images-ligne. Le choix de ces deux paramètres peut s'effectuer expérimentalement et leur réglage se fait alors de manière intuitive et rapide. Certaines cartes d'acquisitions permettent également d'éviter automatiquement le phénomène du décalage spatial par l'utilisation de techniques d'interpolation.

Pour montrer ce phénomène dans un cas pratique, nous avons acquis avec une caméra linéaire couleur l'image d'une feuille blanche sur laquelle du texte avait été préalablement imprimé avec une encre noire. La figure 1.22 est constituée de deux images. L'image 1.22(a) montre l'acquisition effectuée sans le recalage spatial, c'est-à-dire sans que les lignes acquises par les capteurs R, V et B n'aient été recombinées entre elles, tandis que l'image 1.22(b) montre l'acquisition effectuée lorsque les lignes acquises par les capteurs R, V et B ont été recombinées entre elles. On peut constater sur l'image 1.22(a) la présence de fausses couleurs au niveau des transitions noir/blanc dans la direction longitudinale, qui ont disparu dans l'image 1.22(b).
# DECALAGE SPATIAL

(a) Image initiale.

# DECALAGE SPATIAL

(b) Image corrigée.

Figure 1.22 : images-ligne successives juxtaposées dans une image matricielle.

#### 1.3.4.2 Sensibilité des photo-récepteurs

Les capteurs CCD sont constitués d'une couche de silice. Ces atomes de silicium sont sensibles sur un domaine spectral plus vaste que celui de l'oeil humain et qui comprend les infrarouges. En outre, les filtres placés sur chaque cellule photo-sensible laissent passer les infrarouges. La figure 1.23, qui représente les courbes de sensibilité spectrale des capteurs Kodak KLI - 2113 équipant la caméra linéaire couleur Basler L301 - kc montre que ces capteurs sont sensibles au-delà de 800 nm. On peut aussi remarquer sur cette figure que le capteur rouge est particulièrement sensible dans ce domaine. Ceci a pour conséquence de saturer plus rapidement les éléments photo-sensibles recouverts d'un filtre rouge. Il convient donc d'utiliser un filtre infra-rouge devant l'objectif afin d'éviter une saturation excessive des pixels rouges et de ne pas ajouter des charges provenant du domaine de l'infra-rouge qui ne correspondent pas à la synthèse additive de la vision humaine. Suivant la caméra utilisée, le filtre infra-rouge peut être placé dans la caméra [ATM06]. Dans le cas où le filtre n'est pas incorporé, la documentation de la caméra spécifie le type de filtre préconisé par le constructeur [BAS03].



Figure 1.23 : Courbes de sensibilité spectrale du capteur linéaire couleur Kodak KLI-2113 (données constructeur).

Malgré l'emploi d'un filtre infra-rouge, il est très difficile d'obtenir une image d'un objet blanc où les pixels sont caractérisés par des composantes R, V et B de mêmes valeurs. Pour remédier à ce problème, les caméras linéaires couleur permettent d'ajuster les niveaux de tension de chaque pixel à la sortie du registre à décalage. Cet ajustement des niveaux de tension du pixel et par conséquent des niveaux des composantes couleur une fois discrétisées se fait dans les structures VGC de la figure 1.14. Ces structures permettent de régler le gain ou le décalage appliqué à la tension sortant du registre à décalage et donc d'égaliser les niveaux de tension des composantes couleur. Cette opération s'appelle la balance des blancs.

## **1.3.5** Contraintes physiques

Pour résoudre les problèmes d'apparition de fausses couleurs inhérents au phénomène du décalage spatial, il est nécessaire d'utiliser des images-ligne acquises à des instants différents. Ce choix technologique implique une contrainte de mouvement de la pièce devant la caméra et quatre contraintes d'orientation de la caméra que nous avons regroupées sous le terme de contraintes physiques puisqu'elles sont de nature cinématique. Ces contraintes sont présentées dans les paragraphes suivants.

### **1.3.5.1** Orientation des capteurs de la caméra

Les lignes de visées des capteurs CCD de la caméra doivent être parfaitement perpendiculaires à la direction longitudinale sous peine de faire apparaître des fausses couleurs dans la direction transversale. En effet, si cette perpendicularité n'est pas respectée, les éléments photosensibles situés aux mêmes emplacements sur chacun des trois capteurs CCD n'observent pas exactement la même surface du produit à inspecter. Ce phénomène se traduit par l'apparition de teintes orangées et de teintes bleutées suivant la direction transversale si la caméra observe une surface noire sur un fond blanc.

Afin d'illustrer ce phénomène, supposons que l'on souhaite acquérir l'image d'un objet noir sur un fond blanc. Les figures 1.24(a) et 1.24(b) montrent respectivement les lignes de visée des capteurs R, V et B pour une caméra bien positionnée et pour une caméra dont les directions transversales et longitudinales ne sont pas orthogonales respectivement. Nous nous intéressons aux pixels notés p, q situés aux extrémités des capteurs dans la figure 1.24(a) et ceux notés p' et q' dans la figure 1.24(b). Le tableau 1.24(c) renseigne sur les niveaux des composantes couleur de ces pixels et indique la couleur correspondante à ces trois valeurs. On peut remarquer que les pixels q et p sont noirs lorsque le capteur est parfaitement perpendiculaire à la direction longitudinale, ce qui correspond à la couleur de la pièce inspectée contrairement aux pixels q'et p' dont les couleurs sont respectivement oranges et bleues. Ces couleurs qui n'existent pas sur la scène analysée sont dues au mauvais positionnement du capteur qui n'intègre pas le même élément de surface. Ce phénomène est particulièrement visible lors de transitions dans le sens longitudinale dans l'image.

Pour illustrer ce phénomène dans un cas pratique, nous avons acquis avec une caméra linéaire couleur l'image d'une feuille blanche sur laquelle le caractère "L" avait été préalablement imprimé avec une encre noire (voir figure 1.24(d)). L'expérience a été recommencée en pivotant le capteur de manière à ce que son axe ne soit plus perpendiculaire à la direction longitudinale (voir figure 1.24(e)). Sur cette partie, on peut constater l'apparition de fausses couleurs dans la direction transversale dues à cette mauvaise orientation du capteur, contrairement aux transitions sans fausses couleurs de l'image 1.24(d).



(a) Lignes de visée d'une caméra linéaire couleur perpendiculaire à la direction longitudinale.

Pixel	q	q'	р	p'
Capteur R	0	255	0	0
Capteur V	0	128	0	128
Capteur B	0	0	0	255
Couleur				

(c) niveaux des composantes couleur des pixels  $q, \dot{q}$ , p et p' et leurs couleurs associées acquis au même instant. (d) Images-ligne acquises (e) Images-ligne acquises

(b) Lignes de visée d'une caméra linéaire couleur non perpendiculaire à la direction longitudinale.



(d) Images-ligne acquises(e) Images-ligne acquises successivement avec une<sub>successivement</sub> avec une caméra correctement<sub>caméra</sub> dont les capteurs orientée. ne sont pas perpendiculaires à la direction longitudinale.

Figure 1.24 : Exemple de problème causé par la non perpendicularité entre la direction longitudinale et transversale.

## 1.3.5.2 Déplacement du produit à inspecter

La contrainte physique liée au mouvement de la pièce devant la caméra est une trajectoire rectiligne suivant la direction longitudinale. En effet, si la pièce ne se déplace pas dans la direction perpendiculaire à la direction transversale, les éléments photo-sensibles des capteurs R, V et B n'observent pas la même partie de la pièce ce qui donnera l'apparition de fausses couleurs.

Les figures 1.25 et 1.26 illustrent ce problème pour l'acquisition d'une image-ligne d'une pièce avec respectivement un déplacement rectiligne et un déplacement non rectiligne.







(c) Acquisition de la ligne noire de la pièce à inspecter à l'instant  $t + 2\Delta t$  pour le capteur *R*.



(b) Acquisition de la ligne noire de la pièce à inspecter à l'instant  $t + \Delta t$  pour le capteur *V*.

Pixel	p	q
Capteur R	0	0
Capteur V	0	0
Capteur B	0	0
Couleur recombinée		

(d) Valeurs des composantes couleur des pixels p, q et leurs couleurs correspondantes.

Figure 1.25 : Acquisition d'une ligne noire d'une pièce animée d'un mouvement rectiligne suivant la direction longitudinale par une caméra linéaire couleur.

On peut voir sur la figure 1.25(a) les lignes de visée respectives des capteurs R, V et B d'une caméra linéaire couleur, ainsi que les directions longitudinale et transversale. Une pièce grise avec une bande noire posée sur un convoyeur blanc défile devant cette caméra. Afin d'observer cette bande noire, il faut recombiner les lignes acquises aux instants t,  $t + \Delta t$  et  $t + 2\Delta t$  représentées sur les figures 1.25(a), 1.25(b) et 1.25(c) respectivement. Nous nous intéressons aux pixels p et q de ces figures dont les niveaux de composantes sont :

- $-p^B$ ,  $q^B$  pour la composante bleue (ces valeurs correspondent à celles prises par le capteur *B* à l'instant *t* de la figure1.25(a)),
- $p^V$ ,  $q^V$  pour la composante verte (ces valeurs correspondent à celles prises par le capteur V à l'instant  $t + \Delta t$  de la figure 1.25(b)),

 $-p^R$ ,  $q^R$  pour la composante rouge (ces valeurs correspondent à celles acquises par le capteur *R* à l'instant  $t + \Delta t$  de la figure 1.25(c)).

Le tableau 1.25(d) renseigne sur les niveaux des composantes couleurs des pixels p et q ainsi que sur la couleur associée à ces pixels. Ce tableau montre que la couleur caractérisant les pixels p et q est noire. Cette couleur correspond effectivement aux couleurs observées de la pièce.

La figure 1.26 illustre la même scène dans les mêmes conditions que celles de la figure 1.25 hormis une trajectoire non rectiligne de la pièce à inspecter devant la caméra. Les pixels notés p' et q' correspondent aux mêmes emplacements des pixels sur le capteur couleur que ceux des pixels p et q de la figure précédente.



(a) Acquisition de la ligne noire de la pièce à inspecter à l'instant *t* pour le capteur *B*.



(c) Acquisition de la ligne noire de la pièce à inspecter à l'instant  $t + 2\Delta t$  pour le capteur *R*.



(b) Acquisition de la ligne noire de la pièce à inspecter à l'instant  $t + \Delta t$  pour le capteur V.

Pixel	p'	q'
Capteur R	0	0
Capteur V	0	255
Capteur B	0	0
Couleur recombinée		

(d) Valeurs des composantes couleur des pixels q, d, p et p' et leurs couleurs associées acquis aux mêmes instants.

**Figure 1.26 :** Acquisition d'une ligne noire d'une pièce animée d'un **déplacement non rectiligne** suivant la direction longitudinale par une caméra linéaire couleur.

Le tableau 1.26(d) renseigne sur la valeur des composantes couleurs des pixels p' et q' ainsi que sur la couleur associée à ces pixels. Ce tableau montre que la couleur correspondant au pixel

p' est le noir alors que la couleur correspondante au pixel q' est le vert. Cette fausse couleur verte provient du mouvement non rectiligne de la pièce à inspecter devant la caméra. En effet, les surfaces observées par les éléments photo-sensibles pour le pixel q' correspondent à la pièce à inspecter pour les acquisitions faites aux instants t et  $t + 2\Delta t$  tandis que l'élément de surface observé à l'instant  $t + \Delta t$  correspond à une partie du convoyeur blanc. C'est pour cette raison que la valeur de la composante verte du pixel q' vaut 255 alors que celles des composantes R et B sont nulles.

Ces deux exemples expliquent que la pièce à inspecter doit se déplacer de manière rectiligne devant la caméra sous peine de faire apparaître des fausses couleurs suivant la direction longitudinale. Il faut donc éviter l'apparition de vibrations et de glissements qui pourraient déclencher l'apparition de fausses couleurs dans les images-ligne acquises.

## **1.3.5.3** Orientation des capteurs par rapport à la direction transversale

Pour expliquer les deux autres contraintes d'orientation, on se place dans le cas ou les directions transversales et longitudinales sont orthogonales. On définit alors un repère orthonormé  $(O, \vec{i}, \vec{j}, \vec{k})$  dont l'origine O est le point d'intersection entre l'axe optique de l'objectif et le plan défini par la surface inspectée. Le vecteur  $\vec{i}$  est placé suivant la direction longitudinale,  $\vec{j}$  est placé suivant la direction transversale et le vecteur  $\vec{k}$  est normal à la surface à inspectée. Les figures 1.27(a) et 1.27(b) montrent ce système d'axes sur respectivement une vue de côté et une vue en perspective. Soit  $\theta$  l'angle dans le plan  $(O, \vec{j}, \vec{k})$  entre l'axe optique et le vecteur  $\vec{k}$ .



(a) Vue de côté avec  $\theta \neq 0$ .

(b) Vue 3D avec  $\theta \neq 0$ .

**Figure 1.27 :** Problème d'acquisition engendré avec un angle  $\theta \neq 0$ .

La première contrainte d'orientation de la caméra consiste à obtenir  $\theta = 0^{\circ}$ . Les figures 1.27(a) et 1.27(b) montrent le repère avec le système étudié. Ce système est constitué d'un capteur linéaire couleur, d'un objectif représenté par une lentille mince et d'une pièce à contrôler. La pièce à contrôler est noire sur un fond blanc. L'objet à contrôler des figures 1.27(a) et 1.27(b) n'a pas été tracé en noir pour des raisons de visibilité. Dans ces figures,  $\theta$  est fixé à une valeur différente de 0. On peut remarquer sur les figures que cette rotation provoque des lignes de visée des capteurs *R*, *V* et *B* de longueurs et de largeurs différentes. Ceci a pour première conséquence d'intégrer des quantités de lumière différentes. Analysons les pixels p et q situés aux extrémités de l'image-ligne couleur acquise. Sur la figure 1.27(b), les surfaces observées par ces pixels sont représentées par des "rectangles" délimités en noirs. On remarque ainsi que les surfaces considérées des capteurs V et B sont localisées sur la pièce à inspecter, tandis que la surface du capteur R est située en-dehors de la pièce à analyser et renseigne donc, dans notre cas sur le fond blanc. Le tableau 1.28 renseigne sur la valeur des composantes couleurs des pixels p et q ainsi que de la couleur associée au pixel. Ce tableau montre que la couleur caractérisant les pixels p et q est rouge au lieu d'être noire, du fait que la valeur de la composante couleur rouge atteint la valeur 255.

Pixel	р	q
Capteur R	255	255
Capteur V	0	0
Capteur B	0	0
Couleur		

Figure 1.28 : niveaux des composantes couleur des pixels p et q de la figure 1.27.

Cet exemple montre que pour  $\theta \neq 0$ , des halos de couleur rouge ou bleue suivant la valeur de  $\theta$  apparaissent de part et d'autre de l'objet analysé. Ces halos de couleur rouge ou bleue sont d'autant plus importants que  $\theta$  est grand et qu'on s'éloigne du centre de l'image suivant la direction transversale. De plus, les surfaces observées étant différentes pour les capteurs *R*, *V* et *B*, ces 3 capteurs ne reçoivent pas la même quantité de lumière dans ce cas.

## 1.3.5.4 Orientation des capteurs par rapport à la direction longitudinale

Pour la troisième contrainte d'orientation, on définit l'angle  $\psi$  entre l'axe optique et le vecteur  $\vec{k}$  dans le plan  $(O, \vec{i}, \vec{k})$ . L'angle  $\psi$  correspond à une rotation d'axe  $\vec{j}$  dans le plan  $(O, \vec{i}, \vec{k})$ . Les figures 1.29(a) et 1.29(b) montrent le repère avec le même système étudié que précédem-



**Figure 1.29 :** Problème d'acquisition engendré avec un angle  $\psi \neq 0$ .

ment, où  $\psi \neq 0$ . Ceci revient à étudier le cas où le capteur linéaire ne se situe pas dans un plan

parallèle au plan défini par la surface observée. On peut remarquer sur la figure 1.29(b) que les lignes de visée des capteurs R, V et B sont de largeurs et d'orientations différentes. Analysons les pixels p et q situés aux extrémités de l'image-ligne couleur acquise. On remarque que les surfaces observées par les photo-récepteurs R, V et B correspondant au pixel p sont toutes localisées sur la pièce à inspecter, tandis que seule la surface observée par le photo-récepteur V correspondant au pixel q est localisée sur la pièce à inspecter.

Pixel	р	q
Capteur R	0	255
Capteur V	0	0
Capteur B	0	255
Couleur		

Figure 1.30 : niveaux des composantes couleur des pixels p et q de la figure 1.29.

Le tableau 1.30 renseigne sur les niveaux des composantes couleurs des pixels p et q ainsi que sur la couleur associée à ces pixels. Ce tableau montre que la couleur caractérisant le pixel p est noire tandis que celle caractérisant le pixel q est magenta au lieu d'être noire.

Cet exemple montre que pour  $\psi \neq 0$ , des halos de couleur rouge ou bleu peuvent apparaître dans l'image suivant que l'article à analyser arrive devant la caméra linéaire couleur de part et d'autre de l'objet analysé. Ces halos de couleur rouge ou bleu sont visibles suivant la direction longitudinale. Cet effet sera amplifié sur des surfaces non planes, comme des objets avec un axe de révolution. Afin de réduire ces phénomènes, on peut aussi utiliser un objectif télécentrique qui transforme la direction des rayons lumineux de sorte qu'ils soient parallèles entre eux. Cependant, l'emploi d'un objectif télécentrique nécessite une quantité de lumière très importante et le prix de ce type d'objectif est très élevé.

## 1.3.6 Contraintes optiques des caméras linéaires couleur

Les contraintes optiques sont liées à l'utilisation des objectifs. Il existe sur le marché une grande variété d'objectifs de qualités et de prix très variés. Pour certaines applications comme la vidéo surveillance, la qualité de l'optique n'est pas essentielle contrairement aux applications de vision industrielle où la qualité de l'objectif peut s'avérer déterminante. Un objectif est constitué d'un ensemble de lentilles permettant de faire converger les rayons de la scène à acquérir sur le capteur de la caméra. La *focale* d'un objectif détermine les dimensions de la scène vue par le capteur. Appelons f la distance focale de l'objectif, h la longueur de la ligne observée (en mètres), C la longueur du capteur suivant la direction longitudinale (en mètres) et D la distance séparant le centre de la lentille mince de la ligne de visée (en mètres). En considérant l'objectif comme une lentille mince, la focale peut-être calculée par la formule 1.3 :

$$f = \frac{C \times D}{h}.\tag{1.3}$$

Les focales normalisées les plus courantes sont en mm : 3.5, 6.5, 8, 12, 16, 25, 35, 50 et 75 [IMA05b]. Une distance focale courte entraîne un grand angle de vue<sup>5</sup>, on parle alors d'objectif grand angle, tandis qu'un focale longue entraîne un petit angle de vue, on parle alors de téléobjectif. Plus la distance focale est courte, plus les phénomènes de distorsion de l'image sont importants [Mah05]. Le système optique inclus dans l'objectif présente des défauts plus ou moins prononcés suivant la qualité de l'objectif. Les défauts les plus contraignants pour l'acquisition d'images à partir d'une caméra linéaire couleur sont :

- une distribution non uniforme de l'intensité lumineuse sur le capteur pour des éléments de surface identiques. Les niveaux sont plus faibles pour des pixels situés à la périphérie du capteur que pour les pixels situés au centre du capteur. Ce phénomène est du au problème de *vignettage* et à la *loi cosinus puissance* 4. Le problème de vignettage est provoqué par le blocage des rayons lumineux par les parties mécaniques de l'objectif. Le deuxième défaut optique appelé loi de cosinus puissance 4 décroît l'éclairement proportionnellement au cosinus puissance 4 de l'angle de champ. Ce phénomène est d'autant plus important que la taille du capteur est grande et que le diamètre de l'objectif est petit.
- les aberrations chromatiques. Celles-ci résultent de réfractions différentes pour chacune des longueurs d'onde de la lumière lorsqu'elle rencontre les lentilles de l'objectif. Les aberrations chromatiques peuvent être réduites par l'emploi d'objectifs chromatiquement corrigés.

## **1.3.7** Synthèse sur les caméras linéaires couleur

Nous nous sommes intéressés dans cette partie aux caméras linéaires couleur utilisant trois rangées de capteurs R, V et B montées en parallèle, car cette technologie est la plus utilisée pour des applications industrielles. Cependant, l'emploi de ce type de caméra requiert certaines spécificités sous peine de faire apparaître des fausses couleurs dans les images-ligne acquises. Nous avons regroupé ces contraintes en trois familles (technologiques, physiques et optiques) suivant leurs origines. Cette partie permet ainsi, à un utilisateur souhaitant employer une caméra linéaire couleur, de corriger les problèmes spécifiques liés à l'acquisition d'image par ce type de technologie.

Une caméra linéaire couleur peut être reliée à une carte d'acquisition afin de synchroniser les acquisitions d'images-ligne. Nous allons donc détailler dans le paragraphe suivant le principe de fonctionnement des cartes d'acquisition et les possibilités qu'elles offrent pour les caméras linéaires couleur.

## 1.4 Les cartes d'acquisitions linéaires couleur

Le rôle premier d'une carte d'acquisition (ou frame grabber en anglais) est de transférer l'information fournie par la caméra dans une mémoire image accessible par l'utilisateur. Pour

<sup>&</sup>lt;sup>5</sup>L'angle de vue est aussi appelé angle de champ.

cela, une carte d'acquisition permet de définir le nombre d'images-ligne dans l'image matricielle utilisée ensuite par l'unité de traitement et le système de vision. Une carte d'acquisition permet aussi de dialoguer avec différents éléments d'un système de vision. Par exemple, une carte d'acquisition peut, activer certaines entrées automates si la pièce est considérée comme rebut ou synchroniser les acquisitions d'images-ligne d'une caméra linéaire couleur avec le signal provenant d'un codeur incrémental ou d'un capteur de présence. Ces signaux qui donnent l'ordre à la caméra de faire une acquisition sont appelés *signaux de déclenchement* et l'organe donnant naissance à ce signal est appelé *système de déclenchement*.

Conceptuellement, il n'y a pas de différences entre les cartes d'acquisition pour les caméras linéaires monochromes par rapport à celles opérant en couleur. En effet, l'information reçue par le capteur provient toujours de la conversion de l'énergie lumineuse réceptionnée par le capteur en un signal électrique aussi bien pour une caméra couleur que pour une caméra monochrome. Ainsi, il existe des cartes d'acquisition qui permettent de dialoguer avec trois caméras linéaires monochrome et/ou de commander une caméra linéaire couleur. Finalement, les différences entre une carte d'acquisition pour caméras linéaires monochromes ou couleur sont :

- le nombre de caméras que l'on peut connecter à la carte d'acquisition,
- les traitements bas niveaux appliqués aux images-ligne appliqués par des composants FPGA (Field Programmable Gate Array).

Afin de préserver la résolution longitudinale du produit à inspecter, la carte d'acquisition dialogue avec le système de déclenchement des acquisitions. Ainsi, une carte d'acquisition peut être utilisée suivant plusieurs modes classiques de fonctionnement que nous détaillerons. Puis, nous introduirons un mode de fonctionnement original appelé ADR (Advanced Downweb Resampling). Une fois que l'image-ligne acquise est sélectionnée, celle-ci peut subir différents traitements effectués par la carte d'acquisition. Nous présenterons donc, les différents traitements embarqués que peuvent effectuer les cartes à partir des composants FPGA dont elles disposent. Puis, l'image-ligne est recopiée dans une mémoire image matricielle. Nous présenterons alors trois méthodes possibles pour remplir cette mémoire image. Enfin, à cause des progrès technologiques réalisés aux niveaux des caméras, des ordinateurs et des réseaux informatiques, il nous semble intéressant de discuter du rôle des cartes d'acquisition en vision industrielle dans les années à venir.

## 1.4.1 Modes de fonctionnement des cartes d'acquisition

Les cartes d'acquisition permettent soit d'insérer toutes les images-ligne acquises par la caméra dans une mémoire image, soit de sélectionner certaines d'entre elles. Les images-ligne sélectionnées sont extraites du flux d'images-ligne envoyées à la carte d'acquisition en fonction du signal de déclenchement. Le mode de fonctionnement de la carte d'acquisition dépend par conséquent de la provenance du signal de déclenchement. Ce signal peut provenir de la caméra, de la carte d'acquisition ou d'un système de déclenchement externe (généralement un codeur incrémental). Ainsi, il existe quatre types de signaux de déclenchement possibles suivant que :

- la caméra génère son propre signal de déclenchement en mode free running,
- la carte génère un signal de déclenchement interne de fréquence réglable par l'utilisateur,
- la carte génère un signal de déclenchement identique au signal du codeur incrémental auquel elle est reliée,
- la carte génère un signal de déclenchement ré-échantillonné en fonction du signal provenant du codeur incrémental auquel elle est reliée. Ce ré-échantillonnage permet d'opérer à des fréquences d'acquisition différentes à celles délivrées par le codeur incrémental.

## **1.4.2** Mode ADR (Advanced Downweb Resampling)

La société Euresys a développé récemment le mode ADR [Eur05a]. Dans ce mode, les images-ligne sont acquises par la caméra à la fréquence maximale. La carte d'acquisition combine les images-ligne acquises entre deux fronts montant du signal de déclenchement pour fournir l'image-ligne à ranger dans la mémoire image. Ce mode de fonctionnement permet également de travailler avec une période d'acquisition de la caméra inférieure à la période du codeur incrémental en utilisant les images-ligne prises à des instants antérieurs. La figure 1.31 illustre le principe de fonctionnement du mode d'acquisition ADR. On peut y remarquer que les événements du signal de déclenchement sont de plus en plus rapides. Par conséquent, de moins en moins de lignes acquises par la caméra en mode free-running sont combinées entre-elles, jusqu'à ce qu'une même acquisition soit utilisée pour fournir plusieurs images-ligne d'une mémoire image. On peut aussi remarquer sur la figure 1.31, que toutes les images acquises entre deux fronts montants du système de déclenchement sont prises en compte. Ainsi, toutes les parties de l'objet en défilement sont utilisées pour obtenir l'image recombinée.



Figure 1.31 : Principe de fonctionnement du mode d'acquisition ADR.

## 1.4.3 Traitements embarqués sur les cartes d'acquisition

Certaines cartes d'acquisitions permettent de faire des traitements embarqués et par conséquent de gagner du temps CPU pour les autres traitements. Ces traitements embarqués restent encore sommaires du fait que les composants FPGA sont fabriqués à partir de portes logiques de base (ET, OU, ...) [Méc05]. Certaines marques de cartes d'acquisition permettent une programmation personnalisée des composants FPGA en fonction de l'application tandis que d'autres marques ne permettent que des composants FPGA programmés en série pour les traitements suivants :

- gestion du phénomène de décalage spatial,
- transformation des niveaux par une look-up table (LUT),
- correction d'éclairage,
- balance des blancs.

Certaines cartes d'acquisition permettent de gérer le phénomène de décalage spatial expliqué dans le paragraphe 1.3.4 en utilisant des techniques d'interpolation.

Les look-up table ou LUT consistent à modifier la valeur d'une composante couleur d'un pixel en utilisant une fonction mathématique dans le but de modifier la dynamique d'une image. Les LUT sont des outils de traitement d'images issues de l'analyse d'images en niveaux de gris et s'appliquent de manière indépendante sur chaque valeur R, V et B des pixels.

La correction d'éclairage permet de compenser la non-uniformité de l'éclairage, les défauts liés à l'objectif et les différences des sensibilités des photo-récepteurs de l'image-ligne pixel par pixel. La correction d'éclairage consiste à appliquer un gain et un décalage sur chacune des valeurs R, V et B des pixels d'une image-ligne acquise. Les valeurs du gain et du décalage sont déterminées à partir d'une image-ligne de référence noire et d'une image-ligne de référence blanche.

L'acquisition d'images-ligne d'un objet blanc de référence fournit des niveaux de couleur R, V et B différents pour chaque pixel. Ces différences de niveaux de couleur sont principalement dues au système d'éclairage et à la sensibilité des photo-récepteurs. La balance des blancs consiste à égaliser les niveaux de couleur R, V et B des pixels entre les composantes couleur à partir d'un blanc de référence. L'opération consiste à appliquer un gain pour chaque composante couleur des pixels déterminé à partir de l'acquisition d'images-ligne d'un blanc de référence. L'acquisition d'une image-ligne noire permet de mesurer certains bruits présents dans les photo-récepteurs des capteurs R, V et B comme le bruit thermique afin de le soustraire pour les acquisitions ultérieures. L'acquisition d'une image-ligne blanche permet de rendre rectiligne le profil de ligne défini par les niveaux des composantes couleur R, V et B le long des capteurs couleur.

## 1.4.4 Remplissage de la mémoire image

Il existe trois méthodes différentes permettant de stocker les images-ligne acquises par la caméra dans une mémoire image matricielle de dimension définie par l'utilisateur :

- La première méthode consiste à ranger successivement les acquisitions d'images-ligne effectuées dans une mémoire image matricielle tant que celle-ci n'est pas remplie. Cette méthode est utilisée pour faire l'acquisition d'objets de taille fixe et connue.
- La deuxième méthode consiste à remplir plusieurs mémoires images matricielle jusqu'à un signal d'arrêt. Ce mode permet de faire l'acquisition d'objets de grandes tailles ou d'objets de tailles variables.
- La troisième méthode nécessite de définir un nombre de mémoires images matricielles.
   Les acquisitions faites par la caméra sont recopiées dans ces mémoires images matricielles les unes après les autres jusqu'au signal de fin d'acquisition. Si la dernière mémoire image matricielle est remplie et que le signal de fin d'acquisition n'a toujours pas eu lieu, alors la première mémoire image matricielle est remplie est remplie est remplie par les nouvelles images-ligne.
   Cette méthode est utilisée pour faire des acquisitions en continu.

## 1.4.5 Avenir des cartes d'acquisition

Actuellement, les sociétés de conception de dispositif de vision industrielle accentuent leurs efforts sur le développement de caméras intelligentes<sup>6</sup> [Wai03]. Ce type de caméra consiste à introduire l'unité de traitement dans le boîtier de la caméra. Ainsi, la caméra n'a plus besoin d'être reliée à un ordinateur pour effectuer le contrôle d'une pièce puisqu'elle gère elle-même les traitements à effectuer sur l'image acquise et qu'elle peut communiquer avec les autres éléments du système de vision via une carte d'entrée/sortie. Avec cette technologie, l'emploi d'une carte d'acquisition n'est plus nécessaire puisque la caméra n'est plus reliée à un ordinateur. De plus, l'augmentation des débits admissibles pour les transferts de données avec l'emploi du standard Gigabit-Ethernet dont la bande-passante atteint 1 GBit/sec et d'une autre version en préparation avec une bande passante atteignant 10 GBit/sec, permettent un flot de données suffisamment rapide pour des applications de vision en temps réel [AP06].

On peut donc supposer une disparition des cartes d'acquisition d'image dans les années à venir. Les professionnels du marché de la vision pensent que les cartes d'acquisition trouveront leur place uniquement dans les applications où l'on a besoin de synchroniser plusieurs signaux ainsi que pour les applications nécessitant de traiter des flots d'images-ligne à fréquence élevée, ainsi que pour applications nécessitant des caméras analogiques de plus petites tailles [Zan05].

<sup>&</sup>lt;sup>6</sup>aussi appelées caméras intégrées.

## 1.5 Synthèse

Dans ce chapitre, nous avons décrit les trois éléments d'un poste d'acquisition pour le contrôle qualité d'une pièce à inspecter par caméra linéaire couleur suivants :

- le système d'éclairage,
- la caméra linéaire,
- la carte d'acquisition.

Après avoir posé quelques définitions importantes relatives aux caméras linéaires, nous nous sommes intéressés au système d'éclairage pour le contrôle d'aspect par caméra linéaire couleur. Les trois points suivants ont été abordés :

- l'interaction de la lumière avec la matière et plus particulièrement les différents types de réflexion les plus courantes.
- les techniques et géométries d'éclairage pour le contrôle d'inspection par caméra linéaire couleur. Ces techniques consistent à obtenir une lumière émise sur la ligne de visée sans zones d'ombre ni reflets et la plus homogène possible.
- la source lumineuse. Elle doit être une source de lumière blanche qui permet d'exciter les capteurs photo-sensibles *R*, *V* et *B* de la caméra dans les mêmes proportions sous peine d'amplifier de manière critique le bruit présent dans une image-ligne lors d'une opération de correction d'éclairage et/ou de balance des blancs. D'autre part, afin de restituer toutes les couleurs à contrôler par synthèse additive, il est nécessaire d'utiliser une lumière qui émet de l'énergie dans tout le domaine du visible.

Nous nous sommes ensuite particulièrement intéressé aux caméras linéaires couleur utilisant trois rangées de capteurs R, V et B montées en parallèle, car cette technologie est la plus utilisée pour des applications industrielles. Cependant, l'emploi de ce type de caméra nécessite le respect de certaines règles sous peine de faire apparaître des fausses couleurs sur l'image-ligne acquise. Nous avons regroupé ces règles en trois familles (technologiques, physiques, optiques) suivant leurs origines. Cette partie permet ainsi, à un utilisateur souhaitant employer une caméra linéaire couleur, de corriger les problèmes spécifiques liés à l'acquisition d'images par ce type de technologie.

Enfin, nous avons présenté les cartes d'acquisition dont le but est de transférer l'information fournie par la caméra dans une mémoire image matricielle accessible par l'utilisateur. Une carte d'acquisition est l'élément du système d'acquisition permettant de synchroniser les acquisitions d'images-ligne avec un codeur incrémental afin de préserver les résolutions transversales et longitudinales de l'image de l'article à inspecter. Nous avons détaillé pour cela les différents modes de fonctionnement d'une carte d'acquisition pour effectuer ce travail.

Ce chapitre a permis ainsi d'appréhender l'acquisition d'une image représentant un article à inspecter. Les mémoires images matricielles couleur accessibles par l'utilisateur sont codées dans l'espace (R, V, B) puisqu'elles sont issues d'une caméra linéaire couleur dont les éléments photo-sensibles sont recouverts de filtres rouge, vert et bleu. Cependant, il nous semble restrictif de considérer une image couleur comme un ensemble de données multidimensionnelles. En effet, la couleur est une notion très complexe qui obéit à des lois physiques, physiologiques et psychologiques particulières. C'est pour cette raison que le chapitre suivant présente les espaces couleur couramment utilisés en traitement d'images couleur qui respectent ces différents lois ainsi que les principales méthodes d'images qui seront nécessaire à notre application industrielle.

## **Chapitre 2**

## **Représentation et segmentation des images numériques couleur**

Dans le chapitre précédent, nous avons introduit les étapes nécessaires à l'acquisition d'une image couleur par une caméra linéaire. L'image acquise est codée dans l'espace couleur (R, V, B)propre à la caméra utilisée mais il existe de nombreux autres espaces couleur avec des caractéristiques différentes qui permettent d'exploiter de manière pertinente l'information provenant de l'image acquise en vue d'un traitement ultérieur. Le résultat de ce traitement dépend ainsi de l'espace couleur utilisé. L'idée originale des travaux présentés consiste en la sélection automatique des espaces couleur les mieux adaptés à la segmentation d'images couleur.

Pour cette raison, ce chapitre est composé de deux parties différentes. La première a pour objectif de présenter les espaces couleur les plus fréquemment utilisés en traitement d'images et la seconde a pour objectif de présenter la segmentation d'une image couleur.

## 2.1 Représentation des images numériques couleur

Dans cette section, nous présenterons les principaux espaces couleur utilisés en traitement d'images. Nous utiliserons pour cela, la classification des espaces couleur proposée par Vandenbroucke [Van00, VM05]. Celle-ci regroupe les espaces couleur en quatre familles suivant leurs caractéristiques :

- Les espaces de primaires,
- Les espaces luminance-chrominance,
- Les espaces perceptuels,
- Les espaces d'axes indépendants.

Chacune de ces familles est détaillée dans les paragraphes suivants.

## 2.1.1 Les espaces de primaires

La perception humaine de la couleur est caractérisée par son aspect tridimensionnel. Les travaux de Young du tout début du XIX<sup>ème</sup> siècle (vers 1802), repris en 1866 par Helmholtz, mettent en évidence que tout stimulus de couleur (c.f paragraphe 1.3.1) peut être reproduit

par le mélange de trois stimuli : le rouge, le vert et le bleu, appelées primaires ou stimuli de référence [You07, Hel66]. Ce principe est également connu sous le nom de synthèse additive des couleurs.

Il existe ainsi de nombreux espaces couleur qui se différencient par le choix des primaires et du blanc de référence. C'est le cas des espaces liés à l'acquisition d'images couleur (R, V, B). En effet, chaque caméra couleur possède ses propres primaires pour reproduire la couleur des pixels d'une image acquise, ce qui définit un espace (R, V, B) propre à chaque caméra. De même pour la transmission des signaux de télévision analogiques, les téléviseurs américains qui répondent à la norme N.T.S.C. (National Television Standards Committee) utilisent les primaires fixées par la F.C.C. (Federal Communications Commission) tandis que les téléviseurs européens qui répondent à la norme allemande P.A.L. (Phase Alternation by Line) ou à la norme française S.E.C.A.M. (SÉquentiel Couleur À Mémoire) utilisent les primaires fixées par l'E.B.U. (European Broadcasting Union ou Union Européenne de Radio-télévision (U.E.R.)). Cependant, l'espace (R, V, B) défini par la C.I.E. (Commission Internationale de l'Éclairage) reste l'espace de référence car les recommandations faites par cette commission servent de base à la plupart des normes. Les primaires utilisées par la C.I.E. définissent un observateur de référence.

A partir de ces trois primaires, il est possible de leur faire correspondre trois vecteurs orthonormés  $\vec{R}$ ,  $\vec{V}$  et  $\vec{B}$  afin de former un espace vectoriel d'origine *O*. Dans cet espace, chaque stimulus de couleur peut être représenté par un point *C* de coordonnées ( $R_C, V_C, B_C$ ). Les points correspondant à des stimuli de couleur dont les niveaux de couleur<sup>1</sup> sont positifs, sont contenus dans un cube, connu sous le nom de cube des couleurs (c.f. figure 2.1). L'origine *O* correspond à la couleur noire ( $R_C = V_C = B_C = 0$ ) tandis que le blanc de référence est défini par le mélange unitaire des trois primaires ( $R_C = V_C = B_C = 1$ ).

<sup>&</sup>lt;sup>1</sup>Les niveaux de couleur des trois composantes couleur d'un espace couleur sont aussi appelés composantes trichromatiques.



Figure 2.1 : Cube des couleurs (R, V, B).

La droite passant par ces deux points est appelée axe achromatique puisque les points qui la définissent représentent des nuances de gris allant du noir au blanc.

Pour pallier certains problèmes des espaces (R, V, B) telle que la non reproductibilité de certaines couleurs par synthèse additive, d'autres primaires ont été définies par la C.I.E. Ces primaires artificielles déterminées mathématiquement ont donné naissance à l'espace couleur (X, Y, Z). Cet espace couleur peut-être déduit à partir de n'importe quel espace (R, V, B) par une relation matricielle dont la matrice est appelée *matrice de passage*.

Par exemple, sous l'illuminant *E* qui correspond à une lumière d'énergie constante, la transformation d'une couleur codée dans l'espace couleur (R, V, B) de la C.I.E. vers l'espace (X, Y, Z)est donnée par la relation 2.1 :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 2.7690 & 1.7518 & 1.1300 \\ 1.0000 & 4.5907 & 0.0601 \\ 0.0000 & 0.0565 & 5.5943 \end{pmatrix} . \begin{pmatrix} R \\ V \\ B \end{pmatrix}$$
(2.1)

La luminance est définie de manière très précise par l'AFNOR (c.f. p.57 dans [AFN75]). Cependant, pour un souci de compréhension et de lisibilité du manuscrit nous utiliserons dans cette thèse le terme luminance en dehors de son contexte rigoureux de manière à regrouper différentes notions proches comme luminosité, luminance lumineuse, éclairement, clarté, intensité, etc. Ces différents termes sont particulièrement explicités dans les ouvrages de Kowaliski [Kow90] et de Sève [Sèv96].

Les composantes trichromatiques d'un stimulus de couleur sont liées à sa luminance. Deux stimuli de couleur peuvent ainsi posséder le même caractère chromatique appelé *chrominance*, mais avoir des composantes trichromatiques différentes à cause de leur luminance. Afin d'obtenir des composantes qui ne tiennent compte que de la chrominance, il convient de normaliser

les niveaux des composantes trichromatiques par rapport à la luminance. Ceci est réalisé en divisant chaque composante trichromatique par la somme des trois. Les composantes ainsi obtenues, appelées coordonnées trichromatiques, sont donc normalisées et notées r, v et b. Les coordonnées trichromatiques sont alors définies par la relation 2.2 :

$$\begin{cases} r = \frac{R}{R+V+B} \\ v = \frac{V}{R+V+B} \\ b = \frac{B}{R+V+B} \end{cases}$$
(2.2)

Il existe aussi aussi un espace (X, Y, Z) normalisé noté (x, y, z) que l'on détermine à partir de la relation 2.3 :

$$\begin{cases} x = \frac{X}{X + \frac{Y}{Y + Z}}\\ y = \frac{Y}{X + \frac{Y}{Y + Z}}\\ z = \frac{Z}{X + Y + Z} \end{cases}$$
(2.3)

En remarquant que x+y+z = 1 pour la relation 2.3, on peut en déduire un *diagramme de chromaticité* (x, y) dans lequel les deux composantes x et y suffisent à représenter la chrominance d'une couleur. La figure 2.2 représente le diagramme de chromaticité qui contient la courbe passant par les points correspondant à des stimuli de couleur monochromatiques de longueur d'onde variant de 380 nm à 780 nm. Sur le diagramme de chromaticité de la figure 2.2 ont été



**Figure 2.2 :** Diagramme de chromaticité (x, y).

ajoutées les coordonnées (x, y) des émissions d'un corps noir de planck pour différentes températures de couleur ainsi que la courbe des températures de couleur pour les illuminants D. Sur ces courbes, les points *A*, *C* et  $D_{65}$  correspondent à des illuminants souvent approximés par les sources d'éclairage présentées dans le chapitre précédent. L'illuminant *A* correspond à la lumière émise par un corps noir porté à la température de 2856 K. Il correspond à la source lumineuse produite par une lampe à filament de tungstène de 500 W. L'illuminant *C* correspond à la lumière moyenne du jour avec une température de couleur de 6770 K. Enfin, l'illuminant  $D_{65}$  correspond à la lumière du jour avec une température de couleur de 6500 K.

Cette courbe d'émission d'un corps noir a été superposée à un *triangle des couleurs* correspondant aux couleurs que l'on peut obtenir par synthèse additive à partir de trois primaires. Elle met ainsi en évidence qu'une température de couleur basse correspond à une couleur "chaude" (jaune-rouge) tandis qu'une température de couleur élevée correspond à une couleur "froide" (bleu-violet).

## 2.1.2 Les espaces luminance-chrominance

De nombreux espaces couleur ont pour point commun de posséder une composante de luminance et deux composantes de chrominance. Les espaces couleur ayant de tels caractéristiques ont été classés par Vandenbroucke dans une même famille appelée espace de luminancechrominance [Van00, VM05]. Cette famille est divisée par l'auteur en quatre sous familles que nous détaillerons dans les paragraphes suivants et qui sont intitulées :

- espaces perceptuellement uniformes,
- espaces antagonistes,
- espaces de télévision,
- autres espaces de luminance-chrominane.

### 2.1.2.1 Les espaces perceptuellement uniformes

L'espace couleur (X, Y, Z) permet de pallier certains inconvénients engendrés par l'emploi de couleurs primaires "réelles". Cependant, cet espace couleur n'est pas perceptuellement uniforme. Comme le montre MacAdam [Mac42, Mac43], deux couleurs séparées d'une même distance<sup>2</sup> peuvent perceptiblement être considérées comme une seule et même couleur par un observateur humain ou, au contraire comme deux couleurs différentes pour ce même observateur suivant leur position dans le diagramme de chromaticité (x, y). Aussi, pour remédier à ces inconvénients, la C.I.E. propose les espaces couleur  $(L^*, u^*, v^*)$  et  $(L^*, a^*, b^*)$  que l'on peut calculer à partir des composantes trichromatiques d'une couleur exprimée dans l'espace (X, Y, Z) et de celles d'un blanc de référence noté  $(X_b, Y_b, Z_b)$ . Notons que plusieurs travaux ont montré que ces deux espaces couleur ont été jugés équivalents sur le plan de l'uniformité [Rou98, CM02]. Les composantes trichromatiques de ces deux espaces se déduisent des relations suivantes :

$$L^* = \begin{cases} 116 \times \sqrt[3]{\frac{Y}{Y_b}} - 16 & \text{si} & \frac{Y}{Y_b} > 0,008856, \\ 903, 3 \times \frac{Y}{Y_b} & \text{si} & \frac{Y}{Y_b} \le 0,008856. \end{cases}$$
(2.4)

<sup>&</sup>lt;sup>2</sup>selon la distance euclidienne.

Les composantes de chrominance pour l'espace  $(L^*, u^*, v^*)$  sont :

$$u^* = 13 \times L^* \times \left( u' - u'_b \right), \tag{2.5}$$

$$v^* = 13 \times L^* \times (v' - v'_b),$$
 (2.6)

avec: 
$$u' = \frac{4X}{X + 15Y + 3Z},$$
  
 $v' = \frac{9Y}{X + 15Y + 3Z},$  (2.7)

Les composantes de chrominance pour l'espace  $(L^*, a^*, b^*)$  sont :

$$a^* = 500 \times \left( f\left(\frac{X}{X_b}\right) - f\left(\frac{Y}{Y_b}\right) \right), \tag{2.8}$$

$$b^* = 200 \times \left( f\left(\frac{Y}{Y_b}\right) - f\left(\frac{Z}{Z_b}\right) \right), \tag{2.9}$$

avec : 
$$f(x) = \begin{cases} x^{\frac{1}{3}} & \text{si } x > 0.008856 \\ 7.787x + \frac{16}{116} & \text{sinon.} \end{cases}$$
 (2.10)

L'image de la figure 2.3 montre une représentation de l'espace  $(L^*, a^*, b^*)$  sur laquelle on peut remarquer que l'axe de luminance  $L^*$  correspond à une opposition blanc-noir, la composante de chrominance  $a^*$  (ou  $u^*$ ) correspond à une opposition de couleurs vert-rouge et la composante de chrominance  $b^*$  (ou  $v^*$ ) correspond à une opposition de couleurs bleu-jaune. Le principal intérêt de ces modèles est l'emploi d'une métrique euclidienne pour le calcul de distances entre deux couleurs.



Figure 2.3 : Espace couleur  $(L^*, a^*, b^*)$ . Cette image est disponible sur internet à l'adresse : http://tecfa.unige.ch/~lombardf/CPTIC/couleurs/couleur\_ERAG/Pages/ch8p1.htm.

Cependant, on peut s'interroger sur la justesse des résultats pour des calculs exécutés à l'échelle du pixel par des procédures logicielles de traitement d'une image couleur, car cet espace est souvent employé en dehors du contexte rigoureux qui a été défini par la C.I.E.,

à savoir la connaissance des primaires et du blanc de référence qui définissent la matrice de passage vers le système (X, Y, Z) [Van00, Loz98].

Selon X. Zhang et B. A. Wandell, l'espace  $(L^*, a^*, b^*)$  est bien adapté pour comparer les couleurs de deux objets de grandes tailles [ZW96]. Cependant, le calcul de la couleur dans cet espace s'effectue en un point et ignore l'influence de son voisinage. Pour pallier ce problème, l'auteur propose l'espace couleur S-CIELAB. Le calcul de cet espace se décompose en plusieurs étapes :

L'image d'entrée codée dans l'espace couleur (X, Y, Z) est transformée en un plan achromatique d'opposition noir-blanc, un plan vert-rouge et plan jaune-bleu noté (O<sub>1</sub>, O<sub>2</sub>, O<sub>3</sub>) par l'emploi de la matrice 3 × 3 de la relation 2.11.

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} 0.279 & 0.720 & -0.107 \\ -0.449 & 0.290 & -0.077 \\ 0.086 & -0.59 & 0.501 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$
(2.11)

- Chaque plan image codé dans l'espace couleur  $(O_1, O_2, O_3)$  est convolué avec un filtre passe-bas proche de la sensibilité du système visuel humain.
- L'image filtrée est alors transformée à nouveau dans l'espace (X, Y, Z) en utilisant la matrice inverse.
- La dernière étape consiste alors à transformer les couleurs dans l'espace  $(L^*, a^*, b^*)$  à partir des relations, 2.4, 2.8, 2.9 et 2.10.

L'espace S-CIELAB peut donc être considéré comme une phase de pré-traitement à une transformation classique dans l'espace  $(L^*, a^*, b^*)$ .

### 2.1.2.2 Les espaces antagonistes

Cette famille regroupe les espaces couleur basés sur la théorie de Hering. Cette théorie essaie de modéliser le système visuel humain en supposant que l'information couleur perçue par les cônes de la rétine humaine est transformée lors de la transmission au cerveau. Cette théorie se base sur les trois types de cônes suivants :

- Les cônes S qui sont sensibles pour des longueurs d'ondes courtes (short) avec un maximum de sensibilité pour la longueur de 419 nm.
- Les cônes *M* qui sont sensibles pour des longueurs d'ondes moyennes (medium) avec un maximum de sensibilité pour la longueur de 530.8 nm.
- Les cônes L qui sont sensibles pour des longueurs d'ondes longues (long) avec un maximum de sensibilité pour la longueur de 558.4 nm.

Suivant la théorie de Hering, les stimuli reçus par les cônes *L*, *M* et *S* sont transformés en un signal achromatique (de luminance) basé sur l'opposition noir-blanc et en deux signaux chromatiques correspondant à un signal vert-rouge et un signal jaune-bleu appelés aussi *axes cardinaux* [GD03]. Ce type d'espaces couleur est par exemple utilisé pour le tatouage d'images afin d'obtenir des marquages robustes et non perceptibles pour l'homme [ES05].

L'espace proposé par Faugeras a été développé pour être proche d'un espace visuel humain et a été utilisé pour le traitement d'images par l'auteur en 1979 [Fau76, Fau79]. Cet espace couleur se détermine à partir des valeurs des cônes L, M et S. Pour déterminer la couleur dans l'espace (L, M, S) à partir des coordonnées de l'espace (R, V, B) sous l'illuminant  $D_{65}$ , il utilise la matrice de passage de l'équation 2.12 :

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = \begin{pmatrix} 0.2457 & 0.6840 & 0.0703 \\ 0.1101 & 0.7625 & 0.1273 \\ 0.132 & 0.0842 & 0.9026 \end{pmatrix} \cdot \begin{pmatrix} R \\ V \\ B \end{pmatrix}$$
(2.12)

Une fois la couleur exprimée dans l'espace (L, M, S), l'auteur propose d'utiliser un système d'équations afin de déterminer les composantes  $A, C_1$  et  $C_2$ . Ce système d'équations est conditionné par les constantes  $a, u_1, u_2, \alpha, \beta$  et  $\gamma$  et utilise la fonction logarithme qui traduit la non-linéarité des réponses des cônes à un stimulus. Cet espace couleur est ainsi calculé à partir des équations 2.13 :

$$\begin{cases} A = a \left( \alpha \log \left( L \right) \beta \log \left( M \right) + \gamma \log \left( S \right) \right) \\ C_1 = u_1 \left( \log \left( L \right) - \log \left( M \right) \right) \\ C_2 = u_2 \left( \log \left( L \right) - \log \left( S \right) \right) \end{cases}$$
(2.13)

Ainsi, pour coder la couleur dans l'espace de Faugeras  $(A, C_1, C_2)$ , les constantes  $\alpha$ ,  $\beta$  et  $\gamma$  ont été ajustées pour satisfaire au mieux à la fonction de sensibilité spectrale relative de l'oeil. Les valeurs de ces constantes sont :  $\alpha = 0.612$ ,  $\beta = 0.369$  et  $\gamma = 0.019$ . Les constantes a,  $u_1$  et  $u_2$ ont été calculées de manière à rendre l'espace le plus perceptuellement uniforme possible. Les valeurs de ces constantes sont : a = 22.6,  $u_1 = 64$  et  $u_2 = 10$ .

On trouve également dans cette famille l'espace de Garbay [Gar79] dont les composantes issues de simplifications des travaux de Faugeras [Fau76] se déterminent par les relations 2.14 :

$$\begin{cases}
A = \frac{1}{3} \times (\log(R) + \log(V) + \log(B)) \\
C_1 = \frac{\sqrt{3}}{2} \times (\log(R) - \log(V)) \\
C_2 = \log(B) - \frac{\log(R) + \log(V)}{2}
\end{cases}$$
(2.14)

Certains auteurs préfèrent considérer la réponse des récepteurs visuels humain comme linéaire ce qui permet ainsi de supprimer les logarithmes et donc de passer directement d'un espace de primaire (R, V, B) à un espace couleur antagoniste [SB91]. Cet espace couleur, noté (wb, rg, by) et appelé espace de Ballard, se calcule à partir d'un espace de primaires en utilisant les relations 2.15 :

$$\begin{cases} wb = \frac{1}{3} \times (R + V + B) \\ rg = \frac{\sqrt{3}}{2} \times (R - V) \\ by = B - \frac{R + V}{2} \end{cases}$$
(2.15)

J.-B. Braquelaire et L. Brun déterminent l'espace  $(H_1, H_2, H_3)$  à partir du modèle de Faugeras [BB97]. Cet espace est utilisé par les auteurs pour améliorer les résultats de quantification d'images. La composante achromatique correspond à la somme des longueurs d'ondes moyennes et longues reçues par les cônes M et L approximée par  $H_1 = R + V$ . Les deux autres composantes sont des composantes correspondant à un signal d'opposition rouge-vert pour la composante  $H_2$  et un signal d'opposition bleu-jaune pour la composante  $H_3$  comme le suggère la théorie de Hering. Cet espace couleur se détermine par la relation 2.16 :

$$\begin{cases}
H_1 = R + V \\
H_2 = R - V \\
H_3 = B - \frac{R+V}{2}
\end{cases} (2.16)$$

## 2.1.2.3 Les espaces de télévision

Lors du passage de la télévision en noir et blanc à la télévision couleur, il a fallu déterminer un codage qui permet de fournir l'information aussi bien pour les récepteurs noir et blanc que pour les récepteurs couleurs. C'est pourquoi les signaux émis sont décomposés en une information de luminance pour la télévision en noir et blanc et de deux informations de chrominance supplémentaires pour la télévision couleur. Les luminophores des télévisions européennes et américaines étant différents, les primaires notées  $R_F$ ,  $V_F$  et  $B_F$  pour la norme américaine F.C.C. et les primaires notées  $R_E$ ,  $V_E$  et  $B_E$  pour la norme E.B.U. sont différentes. Pour cette raison, il existe un espace couleur pour chacune des deux normes mais il existe des correspondances permettant de passer d'un espace de primaires à l'autre. Ainsi l'espace couleur défini par le standard américain N.T.S.C. est l'espace (Y,I,Q) et l'espace couleur défini par le standard européen P.A.L. est l'espace (Y,U,V). Ces espaces sont définis par les relations 2.17 et 2.18 pour respectivement l'espace (Y,I,Q) et l'espace (Y,U,V).

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.30 & 0.59 & 0.11 \\ 0.60 & -0.28 & -0.32 \\ 0.21 & -0.52 & 0.31 \end{pmatrix} \cdot \begin{pmatrix} R_F \\ V_F \\ B_F \end{pmatrix}$$
(2.17)

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.30 & 0.59 & 0.11 \\ -0.15 & -0.29 & 0.44 \\ 0.61 & -0.52 & -0.09 \end{pmatrix} \cdot \begin{pmatrix} R_E \\ V_E \\ B_E \end{pmatrix}$$
(2.18)

## 2.1.2.4 Les autres espaces

En étudiant les propriétés de différents espaces luminance-chrominance, Carron passe de l'espace (R, V, B) à un espace luminance-chrominance noté  $(Y, Ch_1, Ch_2)$  dans lequel l'équation de luminance reste inchangée par rapport à l'espace de Ballard (équations 2.19) tandis que les équations des composantes de chrominance deviennent [Car95, LC99] :

$$\begin{cases} Y = \frac{1}{3} \times (R + V + B) \\ Ch_1 = R - \frac{V + B}{2} \\ Ch_2 = \frac{\sqrt{3}}{2} \times (B - V) \end{cases}$$
(2.19)

Cet espace, que nous appellerons *espace de Carron*, est utilisé par cet auteur dans le but de déduire les composantes d'un système perceptuel. Cette famille d'espaces est présentée au paragraphe suivant. Notons que la composante  $Ch_1$  de l'espace de Carron correspond à une opposition cyan-rouge et la composante  $Ch_2$  à une opposition vert-bleu.

Dans la famille des espaces luminance-chrominance, nous pouvons enfin citer l'utilisation d'un espace de couleur normalisé, noté (I, r, g), et composé d'une information de luminance I et des deux coordonnées trichromatiques r et g [Nev77, OKS80, SA81, Cou91]. De la même façon, la C.I.E. préconise l'utilisation de l'espace (Y, x, y) où Y représente la luminance et x et y qui sont les coordonnées trichromatiques de l'espace (X, Y, Z), représentent la chrominance.

## 2.1.3 Les espaces perceptuels

Les espaces couleur présentés précédemment ne sont pas les mieux adaptés pour représenter la couleur tel que l'observateur humain la perçoit. En effet, nous ne définissons pas la couleur d'un objet en quantité de rouge, de vert et de bleu mais nous utilisons des termes plus subjectifs comme la teinte, l'intensité et la saturation. Aussi, la première classification de la couleur proposée par A. H. Munsell en 1898 et redéfinie par la Optical Society of America en 1942 utilise ces trois entités.

L'intensité représente la quantité d'énergie d'un stimulus de couleur. Elle correspond donc à une sensation traduite par des vocables comme clair, foncé, lumineux, sombre et caractérise le niveau lumineux d'un stimulus de couleur. Le concept d'intensité a un sens assez large et beaucoup de grandeurs physiques permettent de la quantifier (luminosité, éclairement, luminance, clarté, ...).

La teinte correspond aux dénominations des couleurs telles que rouge, vert, bleu, jaune, ... Elle correspond à l'*intervalle de longueur d'onde dominant* d'un stimulus de couleur, c'est-àdire l'intervalle des longueurs d'ondes pour lequel l'énergie correspondante est la plus élevée.

La saturation correspond au rapport entre l'énergie correspondant à la longueur d'onde dominante et l'énergie correspondant aux autres longueurs d'ondes d'un stimulus de couleur indépendamment de leur quantité. Elle se distingue du *chroma* qui est aussi une grandeur estimant le niveau de coloration d'une teinte mais qui dépend de la quantité d'énergie correspondant à chaque longueur d'onde. La saturation représente la *pureté* de la couleur perçue comme vive, pâle, terne, ...

Les espaces utilisant ces trois entités sont donc classés dans les espaces perceptuels par Vandenbroucke [Van00, VM05]. De nombreux espaces utilisent ces trois attributs perceptuels ; aussi l'auteur divise ce type d'espaces couleur en deux sous-familles :

- les espaces de coordonnées cylindriques qui correspondent à l'expression en coordonnées cylindriques des espaces couleur appartenant à la famille des espaces de luminance-chrominance,
- les espaces humains de perception de la couleur qui sont déterminés à partir d'un espace de primaires.

### 2.1.3.1 Les espaces de coordonnées cylindriques

Cette famille d'espaces découle des espaces de représentation décomposant l'information couleur en un axe de luminance et un plan de chrominance, par transposition des coordonnées

cartésiennes en coordonnées cylindriques. Si nous considérons un point P représentant une couleur dans un espace luminance-chrominance noté  $(L, Chr_1, Chr_2)$ , ce point est défini par les coordonnées du vecteur  $\overrightarrow{OP}$  dans le repère  $(O, \overrightarrow{L}, \overrightarrow{Chr_1}, \overrightarrow{Chr_2})$ . Soit P', la projection de P sur le plan  $(O, Chr_1, Chr_2)$  suivant l'axe L. Sachant que, par cette construction, les vecteurs  $\overrightarrow{OP}$  et  $\overrightarrow{OP'}$ ont la même norme et la même orientation, il est alors possible de repérer le point P grâce à la norme du vecteur  $\overrightarrow{OP'}$ , à l'angle entre les vecteurs  $\overrightarrow{OP'}$  et  $\overrightarrow{Chr_1}$  et à sa coordonnée suivant l'axe L. Les trois composantes qui en découlent sont notées respectivement L, C et H et forment un espace de coordonnées cylindriques qui est noté (L, C, H) (voir figure 2.4).



Figure 2.4 : espace de coordonnées cylindriques.

La première composante d'un espace de coordonnées cylindriques représente l'information de luminance *L*. Elle est donc identique à la première composante de l'espace luminance-chrominance correspondant.

La norme du vecteur  $\overrightarrow{OP'}$  représente le chroma C qui est défini par l'équation :

$$C = \|\overrightarrow{OP'}\| = \sqrt{Chr_1^2 + Chr_2^2}.$$
 (2.20)

Le chroma correspond à la distance du point P à l'axe de luminance. Des points ayant même chroma se situent sur un cylindre centré sur l'axe des luminances que nous appelons *cylindre d'iso-chroma* sur la figure 2.4. Les projections de ces points sur le plan de chrominance définissent le *cercle d'iso-chroma*.

L'angle du vecteur  $\overrightarrow{OP'}$  avec le vecteur  $\overrightarrow{Chr_1}$  représente l'information de teinte, *H*. Afin d'obtenir des valeurs de *H* comprises dans l'intervalle  $[0, 2\pi]$ , son évaluation doit satisfaire aux

conditions de l'équation 2.21.

$$H = (\overrightarrow{Chr_1}, \overrightarrow{OP'}) \text{ avec} \begin{cases} \text{si } Chr_1 > 0 \text{ et } Chr_2 \ge 0 \text{ alors } 0 \le H < \frac{\pi}{2}, \\ \text{si } Chr_1 \le 0 \text{ et } Chr_2 > 0 \text{ alors } \frac{\pi}{2} \le H < \pi, \\ \text{si } Chr_1 < 0 \text{ et } Chr_2 \le 0 \text{ alors } \pi \le H < \frac{3\pi}{2}, \\ \text{si } Chr_1 \ge 0 \text{ et } Chr_2 < 0 \text{ alors } \frac{3\pi}{2} \le H < 2\pi. \end{cases}$$
(2.21)

L'utilisation des relations de trigonométrie permet d'évaluer H qui, par exemple, peut être exprimée de la façon suivante :

$$H = \arccos\left(\frac{Chr_1}{C}\right). \tag{2.22}$$

Cette formulation fournit un angle compris dans l'intervalle  $[0, \pi]$ . Pour obtenir un angle H' dans l'intervalle  $[0, 2\pi]$ , il faut réaliser l'opération : si  $Chr_2 < 0$ ,  $H' = 2\pi - H$ .

Il est possible de construire un espace (L,C,H) à partir de n'importe quel espace luminance-chrominance en utilisant les équations précédentes. Par exemple, la CIE définit les composantes de systèmes de coordonnées cylindriques (L,C,H) à partir des systèmes uniformes  $(L^*, u^*, v^*)$  et  $(L^*, a^*, b^*)$  obtenant ainsi respectivement les systèmes  $(L^*_{uv}, C^*_{uv}, h_{uv})$  et  $(L^*_{ab}, C^*_{ab}, h_{ab})$  [CIE86].

## 2.1.3.2 Les espaces de coordonnées perceptuelles

Ces espaces sont directement évalués à partir d'un espace de primaires et représentent la couleur en terme d'intensité (I), de saturation (S) et de teinte (T). Pour ces espaces, les composantes I, S et T sont déterminées à partir de projections effectuées sur les couleurs incluses dans le cube des couleurs (R, V, B) dont les composantes trichromatiques sont définies sur l'intervalle [0, M]. Suivant la projection considérée, les pixels ayant un même niveau d'intensité donneront naissance à une surface que nous appellerons surface d'iso-intensité. Cette surface peut-être de forme triangulaire ou hexagonale par exemple. Ainsi, les principaux espaces (I, S, T) utilisés en traitement d'images sont le modèle triangulaire, le modèle hexagonal et le modèle hexagonal double. Le modèle hexagonal et le modèle hexagonal double sont représentés respectivement sur la figure 2.5(a) et la figure 2.5(b) de la figure 2.5. Ces deux modèles ainsi que le modèle triangulaire sont des cas particuliers du modèle GLHS proposé par H. Levkowitz et G. T. Herman [LH93].



Figure 2.5 : Modèle hexagonal et hexagonal double.

Nous proposons de rappeler dans cette partie les résultats essentiels de ce modèle sans les détailler afin de ne pas alourdir l'exposé. Le lecteur intéressé trouvera de plus amples informations dans [Han02, Van00, Car95, LH93].

L'approche générale du modèle GLHS consiste à choisir une équation d'intensité qui impose une forme géométrique particulière aux surfaces d'iso-intensité. Soit  $\ell(P)$ , la fonction qui calcule l'intensité associée à une couleur représentée par un point *P* de coordonnées (*R*,*V*,*B*) dans le cube des couleurs. La surface d'iso-intensité *l* est définie par l'ensemble des points qui satisfont la relation 2.23 :

$$\{P = (R, V, B) : \ell(P) = l\}$$
(2.23)

Pour le modèle général des GHLS, l'équation d'intensité est définie par la relation 2.24 :

$$\ell(P) = \omega_{\min}.\min(P) + \omega_{\min}.\min(P) + \omega_{\max}.\max(P)$$
(2.24)

Les fonctions min(P), mid(P) et max(P) donnent respectivement la coordonnée minimum, médiane et maximum du point P et  $\omega_{\min}$ ,  $\omega_{\min}$  et  $\omega_{\max}$  sont trois valeurs de pondération fixées par l'utilisateur, avec  $\omega_{\max} > 0$  et  $\omega_{\min} + \omega_{\min} + \omega_{\max} = 1$ .

Les paramètres  $\omega_{\min}$ ,  $\omega_{\min}$  et  $\omega_{\max}$  déterminent la géométrie de la surface d'iso-intensité. Ainsi, les modèles triangulaire, hexagonal simple et hexagonal double sont déterminés à partir des fonctions d'intensité  $\ell(P) = \frac{R+V+B}{3}$ ,  $\ell(P) = \max(R, V, B)$  et  $\ell(P) = \frac{\max(R, V, B) + \min(R, V, B)}{2}$ respectivement. Le tableau 2.1 indique les valeurs  $\omega_{\min}$ ,  $\omega_{\min}$  et  $\omega_{\max}$  pour obtenir la composante d'intensité de chacun de ces modèles à partir de l'espace GLHS.

	$\omega_{\min}$	$\omega_{\rm mid}$	$\omega_{\rm max}$
modèle triangulaire	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
modèle hexagonal simple	0	0	1
modèle hexagonal double	$\frac{1}{2}$	0	$\frac{1}{2}$

**Tableau 2.1 :** Valeurs des paramètres  $\omega_{min}$ ,  $\omega_{mid}$  et  $\omega_{max}$  des espaces triangulaires, hexagonal simple et hexagonal double à partir du modèle général GLHS.

La teinte associée au point *P* correspond à l'orientation du vecteur couleur  $\overrightarrow{OP}$  par rapport à un autre vecteur de référence qui est généralement le vecteur défini par le point associé à la couleur rouge et l'origine *O*. La teinte est par conséquent indépendante du choix de la fonction d'intensité. Il existe plusieurs formulations différentes qui permettent de calculer la teinte de manière exacte. La relation 2.25 est une formulation couramment utilisée pour calculer cette grandeur [Van00].

$$T = \begin{cases} \arccos\left(\frac{\frac{1}{2}[(R-V)+(R-B)]}{\sqrt{(R-V)^2+(R-B)+(V-B)}}\right) & si \quad B \le V, \\ 2\pi - \arccos\left(\frac{\frac{1}{2}[(R-V)+(R-B)]}{\sqrt{(R-V)^2+(R-B)+(V-B)}}\right) & si \quad B > V. \end{cases}$$
(2.25)

D'autres méthodes de calcul de la teinte de manière exacte ou approchée sont présentées dans [Van00, LH93].

La saturation *S* d'une couleur *P* correspond au rapport des longueurs définies par la longueur entre la couleur *P* et son projeté orthogonal sur l'axe achromatique pour le numérateur et la longueur entre l'origine *O* de l'espace (R, V, B) et le projeté othogonal du point *P* sur l'axe achromatique pour le dénominateur. Pour les espaces GLHS, la saturation est alors obtenue par la relation 2.26 :

$$s(P) = \begin{cases} \frac{\ell(P) - \min(P)}{\ell(P)}, \text{ si } \ell(P) \le \ell(q(P)), \\ \frac{\max(P) - \ell(P)}{M - \ell(P)}, \text{ si } \ell(P) > \ell(q(P)). \end{cases}$$
(2.26)

Pour calculer cette relation, il est nécessaire de calculer des variables intermédiaires q(P), k(P) et f(P) définies respectivement par les relations 2.27, 2.28 et 2.29 :

$$q(P) = \begin{cases} (M, f(P)M, 0), & \text{si } k(P) = 0, \\ (1 - f(P)M, M, 0), & \text{si } k(P) = 1, \\ (0, M, f(P)M) & \text{si } k(P) = 2, \\ (0, (1 - f(P))M, M), & \text{si } k(P) = 3, \\ (f(P)M, 0, M), & \text{si } k(P) = 4, \\ (M, 0, (1 - f(P))M), & \text{si } k(P) = 5. \end{cases}$$
(2.27)

et

$$k(P) = \begin{cases} 0, & \text{si} \quad r > v \ge b, \\ 1, & \text{si} \quad v \ge r > b, \\ 2, & \text{si} \quad v > b \ge r, \\ 3, & \text{si} \quad b \ge v > r, \\ 4, & \text{si} \quad b > r \ge v, \\ 5, & \text{si} \quad r \ge b > v, \end{cases}$$
(2.28)

et

$$f(P) = \begin{cases} \frac{\operatorname{mid}(P) - \operatorname{min}(P)}{\operatorname{max}(P) - \operatorname{min}(P)}, & \text{si } k(P) \text{ est pair,} \\ \frac{\operatorname{max}(P) - \operatorname{mid}(P)}{\operatorname{max}(P) - \operatorname{min}(P)}, & \text{si } k(P) \text{ est impair,} \end{cases}$$
(2.29)

Afin de montrer que l'équation de la saturation proposée est équivalente à celles utilisées pour les modèles triangulaire, hexagonal et hexagonal double, nous montrons que l'équation  $\frac{\ell(P) - \min(P)}{\ell(P)}$  si  $\ell(P) \leq \ell(q(P))$  permet de retrouver les équations traditionnelles de la saturation pour ces trois modèles.

Ainsi, pour le modèle triangulaire, avec l'équation d'intensité  $\ell(P) = \frac{R+V+B}{3}$  on a :

$$\frac{\ell(P) - \min(P)}{\ell(P)} = \frac{\frac{R+V+B}{3} - \min(P)}{\frac{R+V+B}{3}}$$
$$= 1 - \frac{3 \times \min(P)}{R+V+B}$$
(2.30)

Pour le modèle hexagonal dont l'équation d'intensité est  $\ell(P) = \max(R, V, B)$ , on a :

$$\frac{\ell(P) - \min(P)}{\ell(P)} = \frac{\max(P) - \min(P)}{\max(P)}$$
(2.31)

Pour le modèle hexagonal double dont l'équation d'intensité est  $\ell(P) = \frac{\max(R,V,B) + \min(R,V,B)}{2}$ :

$$\frac{\ell(P) - \min(P)}{\ell(P)} = \frac{\frac{\max(P) + \min(P)}{2} - \min(P)}{\frac{\max(P) + \min(P)}{2}}$$
$$= \frac{\max(P) - \min(P)}{\max(P) + \min(P)}$$
(2.32)

L'équation 2.32 correspond à l'équation classique de la saturation de l'espace hexagonal double lorsque  $\ell(P) > \frac{\ell_{max}(P)}{2}$  ou  $\ell_{max}(P)$  correspond à la valeur maximale de l'intensité.

## 2.1.4 Les espaces d'axes indépendants

Comme le souligne [OKS80] les composantes de l'espace couleur (R, V, B) sont fortement corrélées puisqu'elles portent une information de luminance. Ainsi, de nombreux auteurs ont tenté de déterminer des espaces de représentation de la couleur dont les composantes sont indépendantes [Van00, VM05]. Les principales techniques d'analyse de données utilisées pour ce problème sont l'analyse en composantes principales (ACP) et l'analyse en composantes indépendantes (ACI).

#### 2.1.4.1 Analyse en composantes principales et espace d'Ohta

L'analyse en composantes principales (aussi appelée transformation de Karhunen-Loeve) a pour but d'analyser un ensemble de données quantitatives représentées dans un espace multidimensionnel afin d'obtenir un espace de dimension éventuellement réduite dont les composantes, appelées *composantes principales* sont décorrélées. En remarquant que les résultats d'une transformation de Karhunen-Loeve sont similaires sur plusieurs images différentes, Y. I. Ohta, T. Kanade et T. Sakai déterminent un espace d'axes indépendants noté (I1, I2, I3) dont les composantes se déterminent par les relations 2.33 [OKS80] :

$$\begin{cases} I1 = \frac{R+V+B}{3} \\ I2 = \frac{R-B}{2} \\ I3 = \frac{-R+2V-B}{4} \end{cases}$$
(2.33)

Cette transformation permet ainsi de décorréler les axes en évitant des temps de traitement importants par l'emploi d'une transformation de Karhunen-Loeve systématique.

Une étude similaire à celle proposée par Ohta a été proposée par I. Philipp et T. Rath dans le cadre d'une application spécifique [PR02]. Ils en déduisent un espace couleur appelé  $(i_1, i_2, i_3)$  proche de l'espace (I1, I2, I3) dont les composantes se déterminent par les relations 2.34 :

$$\begin{cases} i_1 = (0.34 \times R) + (0.33 \times V) + (0.33 \times B) \\ i_2 = (0.07 \times R) + (0.39 \times V) - (0.54 \times B) \\ i_3 = (-0.35 \times R) + (0.51 \times V) - (0.14 \times B) \end{cases}$$
(2.34)

Y. Chen et P. Hao appliquent la transformation de Karhunen-Loeve sur des images converties dans l'espace  $(L^*, a^*, b^*)$  [CH04]. Les résultats ainsi obtenus étant similaires pour les images testées, ils définissent un nouvel espace appelé LAR dont les composantes se déterminent par les relations 2.35 :

$$\begin{pmatrix} L\\ A\\ R \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\left(\frac{\pi}{4}\right) & \sin\left(\frac{\pi}{4}\right)\\ 0 & -\sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{pmatrix} \cdot \begin{pmatrix} L^*\\ a^*\\ b^* \end{pmatrix}$$
(2.35)

Il est intéressant de remarquer que cet espace conserve l'uniformité de l'espace  $(L^*, a^*, b^*)$ puisque la transformation effectuée correspond à une rotation d'angle  $\frac{\pi}{4}$  par rapport à la composante de luminance. Cette transformation n'apporte par conséquent que peu d'intérêts pour les méthodes de classification utilisées pour la segmentation d'images couleur puisque les formes des nuages des points couleurs calculés à partir d'une image couleur codées dans les espaces couleurs LAR et  $(L^*, a^*, b^*)$  sont identiques.

## 2.1.4.2 Analyse en composantes indépendantes

Dans le cas de l'ACP la première composante déterminée porte le maximum d'informations et les autres composantes lui sont orthogonales. L'ACI contrairement à l'ACP permet de déterminer les axes sur lesquelles l'information est la mieux répartie au sens des moindres carrés et permet ainsi une meilleure représentation des données. En outre, l'ACI serait proche de l'espace visuel des vertébrés en réduisant la redondance des données d'entrée afin d'avoir une représentation interne la plus efficace possible [Bor04].

## 2.1.5 Les espaces hybrides

Il est admis dans la littérature que le choix de l'espace couleur utilisé dans le cadre d'un traitement d'images est crucial car il influe directement sur les résultats de ce traitement. Cependant, ce choix est rendu difficile par la multitude d'espaces couleur existant comme le montre ce chapitre. D'autre part, dans le cadre de la segmentation d'images couleur, il n'existe pas un espace couleur qui soit mieux adapté qu'un autre pour segmenter tous types d'images couleur, comme le souligne H. D. Cheng [CJSW01]. Le choix de l'espace couleur dépend de l'image à analyser d'une part et de l'algorithme utilisé d'autre part.

Partant de ce constat, Vandenbroucke propose de segmenter l'image à partir d'un *espace couleur hybride* qui regroupe des composantes couleur pouvant être issues de différents espace [VMP03]. La segmentation de l'image est alors réalisée par classification des pixels dans l'espace couleur hybride qui est l'espace le plus discriminant vis à vis des classes de pixels en présence.

Contrairement à l'approche que nous proposons dans cette thèse, la construction de l'espace couleur hybride nécessite une phase d'apprentissage supervisé. Durant cette phase, les observations formant l'échantillon d'apprentissage sont sélectionnées interactivement par l'utilisateur sur plusieurs images d'apprentissage et pour chaque classe de pixels à discriminer. A partir de ces observations, les vecteurs d'attributs sont déterminés pour chaque classe de pixels. Les coordonnées de ces vecteurs sont les moyennes des composantes couleur des observations des espaces couleur utilisés. L'auteur met en évidence que d'autres attributs peuvent aussi être utilisés au cours de l'apprentissage comme des attributs de texture par exemple.

La sélection des composantes formant l'espace couleur hybride est effectuée grâce à une procédure de sélection pas à pas associée à l'évaluation d'un critère informationnel. Celui-ci est basé sur des mesures de dispersion et de compacité des classes. Cette procédure classique d'analyse discriminante est enrichie par une phase de rejet des composantes corrélées et par un critère d'arrêt permettant de fournir la dimension adéquate de l'espace couleur hybride.

Cette approche a été principalement appliquée à la segmentation de séquences d'images couleur de matchs de football afin de suivre automatiquement des footballeurs en mouvement. Cette application est présentée dans plusieurs publications ainsi que dans un récent ouvrage collectif sur l'imagerie couleur [VMVP04, VMP03, VMP98].

## 2.1.6 Synthèse sur les espace couleurs

La figure 2.6 présente le regroupement des espaces tel qu'il a été proposé par Vandenbroucke [Van00, VM05] et dans ce chapitre. Les quatre rectangles grisés présentent les quatre familles d'espaces couleur :

- les espaces de primaires,
- les espaces de luminance-chrominance,
- les espaces perceptuels,
- les espaces d'axes indépendants.

Les flèches indiquent les transformations possibles entre espaces. Certaines transformations non linéaires ne sont pas réversibles comme le passage d'un espace d'acquisition d'images vers un espace de coordonnées perceptuelles [CJW02, Shi95].



Figure 2.6 : Les familles d'espaces couleur.

L'utilisation de ces espaces doit s'effectuer dans des conditions rigoureuses. Par exemple, pour certains espaces, les primaires et le blanc de référence doivent être connues. Dans le cadre du traitement d'images, la transformation de l'espace (R,V,B) dans un autre espace modifie souvent la dynamique des composantes couleur. Ainsi, l'intervalle de valeurs entières non signées auquel appartiennent les niveaux de chaque composante R, V et B se transforme en un intervalle de valeurs réelles signées ayant des bornes différentes. Afin d'utiliser et comparer des images converties dans différents espaces couleur, Vandenbroucke montre qu'il est nécessaire d'effectuer un codage de la couleur qui respecte les propriétés de chaque espace, notamment pour l'évaluation d'histogrammes [VMP00]. Cet aspect étant souvent négligé par les auteurs, l'auteur propose un codage permettant la comparaison d'images, d'histogrammes ou de distances dans différents espaces couleur tout en préservant les propriétés que possèdent chacun de ces espaces.

Ce chapitre a permis d'introduire différents espaces couleur souvent utilisés en traitement d'images. Cette liste bien que non exhaustive, met en avant la difficulté qu'il peut y avoir à utiliser et choisir un espace plutôt qu'un autre tant le nombre d'espaces couleur est important. Dans notre étude concernant la segmentation d'images couleur, nous utiliserons certains des espaces présentés afin de segmenter une image couleur dans un espace couleur adapté. Aussi, après avoir présenté les principaux espaces couleur, nous présentons la problématique liée à la segmentation d'une image couleur et établissons un état de l'art des méthodes de segmentation.

## 2.2 Segmentation des images numériques couleur

Conceptuellement, la segmentation consiste à extraire de l'image des primitives, soit de type contour, soit de type région. Ces primitives seront exploitées ultérieurement pour effectuer de la reconnaissance de formes, du recalage, de la mise en correspondance, de la compression... La segmentation n'est donc pas un objectif en soi, mais une étape située en amont de dispositifs d'interprétation d'images, de reconstruction tridimensionnelle, de codage, etc.

À la différence du prétraitement, qui transforme l'image couleur en une autre image couleur, la segmentation transforme l'image couleur soit en une image binaire contours/non contours, soit en une image de régions pour laquelle tous les pixels d'une même région sont affectés d'une même étiquette. Il y a donc nécessairement passage de trois composantes à une seule composante. Cette étape peut être vue comme la fusion intervenant, soit au début afin de réduire l'espace couleur à un espace monochrome, soit à la fin du traitement [TFMB04].

Il existe trois approches possibles pour le traitement des images couleur en vue de la segmentation :

- Passage d'une image couleur à une image monochrome ; cela peut se faire, par exemple, par la transformée de Karhunen-Loève ou par une quantification. On a alors à disposition tout l'arsenal des méthodes de segmentation monochrome.
- Approche marginale : les composantes couleur sont traitées séparément, les trois résultats

sont ensuite fusionnés.

Approche vectorielle : chaque pixel est traité comme un vecteur dans l'espace couleur.
 Les primitives sont extraites en tenant compte des trois composantes à la fois.

La littérature foisonne de méthodes de segmentation, mais aucune ne permet d'obtenir des résultats satisfaisants pour toutes les images. Le choix d'une méthode dépend souvent des connaissances *a priori* sur l'image et repose sur les différents aspects suivants :

- Le choix entre extraction des contours ou extraction des régions ;
- La présence ou non de texture, de bruit ;
- Le choix de l'espace couleur et éventuellement d'une métrique dans cet espace ;
- Le but recherché : extraire tous les objets, obtenir des contours localisés ou formant une chaîne fermée de contours, isoler des régions pertinentes.

Nous nous focaliserons uniquement sur les méthodes de segmentation d'images qui utilisent l'extraction de régions car notre contribution intervient sur ce type de méthodes. Les méthodes de segmentation d'images en régions s'attachent à rechercher les zones de l'image dont les couleurs sont homogènes.

Formellement, la segmentation d'une image I définie sur un ensemble de sites S est une partition de S en sous-ensembles disjoints non vides  $\Re_i$  pour i = 1, 2, ..., M, appelés régions contenant des sites connexes au sens d'un prédicat et tels que :

$$S = \bigcup_{i=1}^{M} \Re_i \tag{2.36}$$

La segmentation en régions a pour but de regrouper des pixels connexes ayant des attributs similaires (couleur, texture...), afin de constituer des régions homogènes et cohérentes, ces régions étant dans la mesure du possible liées à un objet ou à une partie d'objet de la scène. Dans les images couleur, les attributs les plus utilisés sont issus de l'information achromatique et l'information chromatique.

Les méthodes de segmentation par construction de régions peuvent être divisées en deux familles suivant que les pixels sont analysés dans le plan image ou dans l'espace couleur. Ainsi, les méthodes de construction de régions par analyse spatiale du plan image parcourent l'image afin de regrouper des pixels connexes respectant le prédicat utilisé pour construire les régions. Les méthodes de segmentation par classification de pixels dans l'espace couleur associent à chaque pixel *P* un point couleur *C*(*P*). Dans le cas de l'espace (*R*,*V*,*B*), les coordonnées de ce point dans l'espace (*R*,*V*,*B*) sont les composantes couleur (*C*<sup>*R*</sup>(*P*),*C*<sup>*V*</sup>(*P*),*C*<sup>*B*</sup>(*P*)) du pixel *P*. Ces méthodes supposent que les pixels appartenant à une même région donnent naissance à un nuage de points couleur dans l'espace couleur. Chacun de ces nuages définit une classe de pixels de l'image. Les pixels sont alors assignés aux différentes classes construites. Une région est ainsi définie comme un ensemble de pixels connexes assignés à une même classe. En effet, plusieurs régions peuvent donner naissance à une même classe et seule une analyse en composante connexe permet de les différencier.

Les deux paragraphes suivants de ce chapitre présentent respectivement les méthodes de
segmentation mettant en oeuvre une analyse spatiale de l'image et celles qui analysent la distribution des points couleurs dans un espace couleur en vue d'effectuer une segmentation par classification de pixels.

Enfin, il est important de noter que quelques méthodes font coopérer les deux approches précédentes.

#### 2.2.1 Segmentation d'images couleurs par analyse spatiale

Comme nous l'avons vu précédemment, de nombreuses méthodes de segmentation existent. Parmi celles-ci, nous pouvons distinguer les méthodes qui analysent directement l'image, privilégiant la disposition spatiale des pixels, de celles basées sur l'analyse de la distribution des points couleurs. Pour toutes ces méthodes, la décision d'associer des pixels afin de construire des régions est binaire ; le prédicat utilisé est donc soit vrai soit faux.

Deux principales approches existent pour reconstruire les régions : l'approche *ascendante* et l'approche *descendante*. En partant d'un pixel, l'approche ascendante consiste à parcourir l'image à partir d'un balayage qui peut être complexe de telle sorte que les pixels voisins qui ont des propriétés couleur proches de celles du pixel courant se regroupent afin de reconstruire les régions au fur et à mesure du balayage. A l'inverse, l'approche descendante se définit comme une succession de divisions spatiales de l'image en blocs ayant des propriétés couleur différentes. Les méthodes de segmentation par croissance et par fusion de régions constituent les approches ascendantes. Nous ne présenterons pas les approches descendantes qui sont très peu exploitées du fait de la complexité des images.

Les méthodes de segmentation par analyse spatiale privilégient essentiellement l'interaction spatiale entre les pixels. Les différences entre les méthodes de ce type proviennent principalement des critères utilisés par les prédicats qui sont par exemple :

- la distance entre les couleurs de deux pixels (généralement la distance euclidienne) [Muk02, Cou91],
- la discontinuité dans le voisinage d'un pixel mesurée par la valeur d'un gradient local [ZL98, FYEA01],
- un critère d'homogénéité couleur de la région en cours de reconstruction [TB97, SC05],
- un critère fondé sur l'analyse des textures couleur [DMS99].

Parmi les approches ascendantes, la segmentation par croissance de régions consiste à explorer l'image à partir de regroupements initiaux de pixels connexes appelés germes, et à faire croître ceux-ci sous une double contrainte d'adjacence et d'homogénéité. Les germes sont sélectionnés de manière déterministe ou aléatoire par pré-traitement de l'image. De manière itérative, de nouveaux pixels sont ajoutés aux germes jusqu'à ce que la région en cours de construction ne respecte plus le ou les prédicats mis en place. La plupart de ces algorithmes parcourent l'image selon un ordre qui peut influencer le résultat. Le parcours le plus fréquent consiste à partir d'un germe en haut à gauche de l'image pour se diriger vers le bas à droite de l'image.

A. Trémeau et N. Borel proposent un algorithme de croissance de régions dont le premier

germe est le pixel en haut à gauche [TB97]. Les pixels voisins du germe sont regroupés à la région en cours de construction s'ils respectent les trois critères suivants : un critère d'homogénéité locale entre chaque pixel voisin et la région, un critère d'homogénéité globale de la région et un nombre minimum de pixels pour construire une région. Si un pixel voisin ne respecte pas l'une de ces conditions, une nouvelle région est constituée à partir de ce pixel et le processus est réitéré. Cependant, la méthode nécessite l'utilisation de trois seuils dont les valeurs sont difficiles à ajuster. De plus, cette méthode a tendance à fournir une sur-segmentation de l'image et nécessite généralement une phase de fusion de régions utilisant d'autres critères.

Afin d'améliorer les résultats, des techniques hybrides ont émergé comme l'approche contour-région. En effet, nous avons vu que la croissance de régions dépend des germes utilisés, et la détection de contours peut naturellement améliorer leur localisation. C'est ainsi que D. Zugaj et V. Lattuati, mais aussi J. Fan et Al., proposent une approche contour-région en détectant les contours de l'image couleur par la recherche des maximums locaux de la réponse au gradient de Di Zenzo [Zen86]. Les germes initiaux sont les régions délimitées par les lignes fermées formées par les pixels contours ainsi détectés [ZL98, FYEA01].

J. Mukherjee utilise les champs de Markov pour limiter la sur-segmentation provoquée par une croissance de régions initiale [Muk02]. En partant du pixel en haut à gauche, l'agglomération des pixels à la région en cours de construction se décide selon la distance Euclidienne maximale entre la couleur du pixel à agréger et celles des pixels de la région en cours de construction. Il résulte de cette pré-segmentation une sur-segmentation de l'image et la présence de pixels non classés car associés à des régions dont le nombre de pixels est insuffisant. Une analyse par modélisation des champs de Markov est alors appliquée afin d'éliminer ces régions.

L. Shafarenko et M. Petrou exploitent une image en niveaux de gris qui représente la réponse d'un gradient couleur pour faire croître les régions par un algorithme de ligne de partage des eaux (LPE) issu de la morphologie mathématique [SP97]. Selon la terminologie géologique, la LPE est définie par la crête que forme la limite entre deux bassins versants [CP95]. Une image en niveaux de gris peut en effet être perçue comme un relief si nous associons le niveau de chaque pixel à une altitude. La ligne de crête est obtenue en "remplissant" chaque bassin versant initialement défini par un minimum local. L'efficacité de la ligne de partage des eaux, en tant qu'outil de segmentation, dépend essentiellement des marqueurs initiaux qui sont ici définis comme les minima locaux de l'évaluation du module d'un gradient couleur. Comme les techniques précédentes, cette approche aboutit souvent à une sur-segmentation. Malgré un filtrage adéquat des minimas locaux, une phase de fusion de régions par l'analyse d'un graphe d'adjacence de régions (GAR) est nécessaire.

L'approche fusion de régions est fondée sur l'analyse d'un GAR qui analyse une image pré-segmentée [TC00]. Ce GAR est une structure composée de nœuds représentant les régions, reliés par des arêtes formalisant l'adjacence entre ces régions. Les régions adjacentes fusionnent si elles respectent des critères d'homogénéité basées sur des distances couleur entre les pixels de

deux régions. Comme pour la croissance de régions, l'ordre du balayage du graphe est important et peut modifier le résultat de la segmentation finale.

Après avoir présenté les principales méthodes de segmentation d'image couleur par analyse spatiale, nous allons maintenant nous intéresser aux méthodes de segmentation par classification de pixels.

#### 2.2.2 Segmentation par classification de pixels

Dans les paragraphes suivants nous présentons essentiellement des méthodes issues de la classification de données multi-dimensionnelles qui sont appliquées à la segmentation d'images couleur. Ces méthodes ne prennent pas en compte la disposition spatiale des pixels et ne considèrent que la distribution des couleurs dans l'espace couleur. Elles construisent les classes de pixels en présence dans l'image et affectent une étiquette à chaque pixel correspondant à sa classe d'appartenance. La formation des régions n'est obtenue qu'après une analyse de la connexité des pixels dans l'image étiquetée.

La construction des classes de pixels peut s'effectuer de deux manières selon que l'algorithme analyse la distribution des points couleur selon une approche métrique ou une approche statistique.

#### 2.2.2.1 Approche métrique

Ces méthodes consistent en une analyse des nuages de points couleur pour construire les classes. La première méthode est issue de l'algorithme des C-means. Cet algorithme est basé sur la minimisation des carrés des distances euclidiennes entre chaque point couleur et le centre de gravité de chacune des classes en présence [Mac67]. La minimisation de ce critère nécessite la recherche exhaustive de toutes les partitions possibles de pixels en classes. L'implantation directe de cette méthode n'est pas envisageable pour la segmentation. L'algorithme des C-means généralement utilisé nécessite la connaissance a priori du nombre  $N_C$  de classes. Après un tirage aléatoire des pixels qui vont chacun définir les centres de gravité initiaux des  $N_C$  classes, l'algorithme consiste, pour chaque itération, à :

- assigner tous les pixels de l'image à la classe la plus proche,

- mettre à jour les centres de gravité des classes ainsi constituées.

L'algorithme est itéré jusqu'à ce qu'un critère basé sur la stabilisation des classes ou des centres de gravité soit atteint.

Afin d'obtenir des résultats indépendants de la position initiale des centres de gravité, T. Uchiyama et M.A. Arbib proposent de segmenter les images couleurs en utilisant un apprentissage compétitif [UA94]. L'algorithme commence avec une seule classe dont les coordonnées du centre de gravité correspondent à la couleur moyenne de tous les pixels de l'image. La méthode consiste à sélectionner aléatoirement un pixel parmi tous les pixels de l'image et à assigner ce pixel à la classe dont le centre de gravité est le plus proche. Cette classe est considérée comme la classe gagnante. Le centre de gravité de la classe gagnante est mis à jour, et un compteur de victoires est incrémenté. La mise à jour des centres des classes fait intervenir un coefficient d'apprentissage dont le rôle est d'assurer leur convergence. Si le compteur de victoires atteint un seuil prédéterminé et que le nombre de classes créées reste inférieur à un nombre de classes maximal prédéterminé, alors une nouvelle classe est créée avec pour centre de gravité la couleur du pixel considéré. Le compteur de victoires de ces deux classes est alors remis à zéro. Cet algorithme est réitéré jusqu'à ce que le nombre de classes ou que le nombre maximal d'itérations soit atteint.

L'algorithme des fuzzy c-means est un algorithme de classification floue qui est une amélioration de la méthode précédente car elle considère que chaque pixel appartient aux différentes classes selon des degrés d'appartenance compris entre 0 et 1 [Bez81]. Il est fondé sur l'optimisation d'un critère quadratique de classification où chaque classe est représentée par son centre de gravité. L'algorithme nécessite de connaître au préalable le nombre de classes et de fixer un paramètre flou qui décrit l'exposant pondérant le degré d'appartenance. Cette méthode construit les classes de telle sorte que la somme des écarts quadratiques intraclasses soit minimale.

Macaire propose une méthode originale qui associe à la fois l'analyse de la distribution des points couleur dans l'espace couleur et l'analyse spatiale des pixels dans l'image couleur [MVP06, MVP04]. La méthode proposée partitionne d'abord l'espace couleur en domainescouleur. Puis elle analyse les propriétés de connexité spatiale et d'homogénéité couleur des sous-ensembles de pixels définis par ces domaines-couleur afin de construire les classes de pixels qui correspondent effectivement aux régions de l'image. Pour sélectionner des domainescouleur pertinents, l'auteur propose une nouvelle mesure, le degré de compacité spatio-colorimétrique qui évalue le degré de confiance que l'on peut placer dans la proposition "le sous-ensemble de pixels défini par le domaine-couleur correspond effectivement à une région de l'image".

#### 2.2.2.2 Approche statistique

L'histogramme d'une image couleur représente une estimation de la probabilité d'apparition des couleurs des pixels de l'image. Un histogramme couleur est implanté par une structure tridimensionnelle (*3D*) composée de cellules caractérisant chaque couleur et contenant chacune le nombre de pixels de la couleur correspondante contenus dans l'image. Quand chaque composante couleur est codée sur 8 bits, l'histogramme couleur contient ainsi 256<sup>3</sup> cellules. Il nécessite une grande place mémoire, ce qui n'a incité jusqu'alors que peu d'auteurs à l'analyser directement [LM00]. Une solution consiste à quantifier l'image couleur différemment en ne considérant que les bits de poids forts, mais cette solution donne de mauvais résultats car elle se résume à une pré-segmentation.

Le principe général des procédures d'analyse d'histogrammes consiste à analyser la distribution dans l'espace couleur des points couleur de manière à mettre en évidence soit des zones à fortes concentrations locales de points couleurs que nous appelons "domaines modaux", soit des zones de l'espace relativement vides de points couleurs qui correspondent aux "vallées" séparant les différents domaines modaux. Sous l'hypothèse généralement admise qu'à chaque domaines modal est associé une classe de pixels, le problème de la détection des classes nécessaires à la segmentation peut-être abordé par celui de la détection des domaines modaux. Ces méthodes aboutissent à la classification des pixels par l'intermédiaire de deux étapes successives : une étape de détection des domaines modaux et une étape de classification des pixels s'appuyant sur les domaines modaux détectés.

Pour cela, il existe deux approches possibles : l'approche marginale et l'approche vectorielle. La première traite les composantes couleurs séparément, et les trois résultats sont ensuite fusionnés. L'approche vectorielle traite chaque pixel comme un vecteur dans l'espace couleur.

#### Approche marginale

Nous venons de voir qu'un domaine modal est défini par une zone de l'espace couleur à forte concentration locale de points couleurs. Les approches marginales analysent la distribution de la projection de ces points sur chacune des composantes couleur. L'analyse marginale consiste donc à analyser indépendamment les trois histogrammes 1D dont chacun correspond à une estimation de la fonction de densité de probabilité des points couleur projetés sur une composante.

Les points couleur des pixels constituant une même région se projettent sur un intervalle le long d'une composante et l'histogramme 1D associé à cette composante va alors présenter un pic sur ce même intervalle. L'analyse de chaque histogramme 1D consiste donc à détecter ces pics, qui peuvent représenter plusieurs régions de couleurs différentes, par l'identification des seuils les délimitant. Les seuils détectés selon chaque composante permettent de partitionner l'espace 3D en différents parallélépipèdes rectangles dans lesquels sont inscrits les différents domaines modaux à détecter. Une analyse de la population des pixels dont les couleurs sont incluses dans ces parallélépipèdes permet de retrouver les domaines modaux.

L'analyse des différents histogrammes 1*D* se fait alors soit de manière itérative afin de détecter au fur et à mesure des itérations les différents parallélépipèdes rectangles dans lesquels sont inscrits les différents domaines modaux associés aux classes, soit en effectuant une analyse hiérarchique des différents histogrammes 1*D* [LM00]. Généralement, ces méthodes supposent que la distribution couleur de chaque classe est Gaussienne. Elles se distinguent notamment par le choix des composantes couleurs utilisées et par la méthode de détection des pics présents dans les histogrammes [Fon01].

R. Ohlander *et al* furent les premiers à étudier la segmentation par approche récursive en prenant en considération neuf composantes issues de trois espaces couleurs différents [OPR78]. Les pics dans chacun des neuf histogrammes sont détectés selon selon des critères hiérarchiques basés sur leur amplitude et leur position. Parmi tous les pics des neuf histogrammes, le pic associé à la plus grande taille de population de pixels est sélectionné. Ces pixels sont assignés à la première classe ainsi construite. Le processus est réitéré sur la nouvelle image ne contenant plus les pixels assignés à cette classe. Cette approche permet de mettre en évidence les différentes classes de pixels au fur et à mesure des itérations. Les critères de sélection des pics sont

très simples et ne prennent pas en compte la répartition spatiale dans le plan image des niveaux inclus dans chaque pic. Ainsi, les pics sélectionnés ne correspondant pas toujours aux régions de l'image, mais plutôt à des groupes de régions. D'autre part, les auteurs n'utilisent qu'une seule composante couleur pour construire une classe. Cette approche a donc tendance à fournir une segmentation grossière de l'image.

Cette approche est reprise par Y.I. Ohta *et al* en montrant que cette méthode est plus efficace quand les composantes couleur utilisées sont décorrélées [OKS80]. Fort de cette conclusion, certains auteurs ont par la suite privilégié les espaces  $(L^*, a^*, b^*)$  et  $(L^*, u^*, v^*)$  pour développer leurs algorithmes [Cel90, Sch93].

S.Tominaga segmente les images couleur par analyse récursive d'histogrammes mono-dimensionnels en adaptant l'espace de représentation de la couleur aux données qui doivent être traitées [Tom92]. A chaque étape, il effectue une transformation de Karhunen Loeve afin de construire les classes. Il analyse l'histogramme du premier axe qui est l'axe le plus discriminant. Si plusieurs pics significatifs de l'histogramme sont détectés, les classes correspondant aux différents pics sont construites, et une nouvelle transformation de Karhunen Loeve est effectuée avec les pixels qui n'appartiennent pas à ces classes. Si un seul pic de l'histogramme du premier axe est détecté, alors c'est l'histogramme du deuxième axe qui est analysé, et ainsi de suite...L'algorithme s'arrête lorsque l'histogramme du troisième axe ne possède plus qu'un seul pic ou lorsqu'il n'existe plus aucun pic significatif parmi les trois histogrammes.

H. D. Cheng et Y. Sun proposent d'utiliser l'histogramme des homogénéités pour détecter les modes [CS00]. Chaque pixel de l'image est caractérisé par l'homogénéité de son intensité, à savoir le rapport entre une mesure locale de dispersion de l'intensité et une mesure de la discontinuité de l'intensité. La mesure locale de dispersion est définie par la différence entre les intensités des pixels appartenant au voisinage du pixel central et la moyenne des intensités des pixels de ce même voisinage. La discontinuité de l'intensité est donnée par la réponse à l'opérateur de Sobel. L'histogramme des valeurs d'homogénéité est alors construit. Dans cet histogramme, rechercher les modes revient à rechercher les régions homogènes en intensité de l'image sans tenir compte de leurs couleurs.

#### Approche vectorielle

L'analyse des histogrammes 1*D* développée précédemment opère de façon récursive afin de restreindre l'influence du recouvrement de la distribution des différents domaines modaux. Cependant, des domaines modaux non discernables par l'analyse d'histogrammes 1*D* peuvent être extraits en analysant simultanément plusieurs composantes couleurs. Il s'agit alors de l'analyse de l'histogramme couleur. La méthode la plus simple pour analyser l'histogramme couleur consiste à effectuer une simple binarisation des valeurs contenues dans les cellules. En effet, les modes sont dans ce cas assimilés aux cellules ayant les plus fortes valeurs et une simple comparaison à un seuil permet leur extraction. L'inconvénient majeur de cette méthode réside dans l'ajustement délicat du seuil.

F. Kurugollu *et al* proposent de décomposer le problème en l'analyse de plusieurs histogrammes 2D, (R,B), (R,V) et (B,V) [KSH00]. Chaque histogramme 2D est alors filtré et K domaines modaux sont détectés par une recherche de maxima locaux des contenus des cellules. Le nombre K de domaines modaux à détecter est fourni par l'utilisateur et une opération de fusion de régions permet finalement d'obtenir l'image segmentée.

La détection des domaines modaux peut aussi être réalisée en utilisant des approches morphologiques comme la ligne de partage des eaux (LPE) ou le filtrage morphologique. L. Shafarenko et al proposent d'appliquer l'algorithme de LPE sur l'histogramme 3D dans l'espace couleur  $(L^*, u^*, v^*)$  [SPK98]. Leur approche consiste à réduire dans un premier temps le bruit d'acquisition par un filtre Gaussien. L'histogramme ainsi filtré est ensuite analysé par un algorithme de LPE afin de détecter les différents domaines modaux en présence. Une approche similaire est proposée par T. Géraud *et al* [GSD01]. L'histogramme 3D est calculé et une fonction logarithmique y est appliquée pour limiter l'incidence de la non-équiprobabilité des classes à détecter. Un filtrage permet ensuite de supprimer les minima locaux de l'histogramme 3D pour ne garder que les marqueurs principaux. Finalement, l'algorithme de LPE est appliqué à cet histogramme filtré où chaque bassin détecté correspond à un domaine modal. Une phase de classification des pixels aux classes associées aux domaines modaux détectés permet d'obtenir l'image segmentée.

S. H. Park *et al* proposent de détecter les domaines modaux associés aux classes dans l'histogramme 3D grâce à des filtres morphologiques [PYL98]. Dans un premier temps, l'histogramme 3D original est filtré par deux filtres Gaussiens d'écart-types différents. La différence des deux histogrammes ainsi obtenue permet de faire ressortir les domaines modaux principaux par une binarisation des cellules dont la population de pixels est significative. Une dilatation adaptative est appliquée pour extraire les modes. Cette dernière opération est adaptative car elle permet aux domaines modaux de s'étendre jusqu'à une certaine limite et selon certaines directions. En effet, si deux modes proches l'un de l'autre se rencontrent durant la dilatation adaptative, le processus est arrêté dans la direction dans laquelle ils tendent à s'étendre. Lorsque les domaines modaux ne peuvent plus être dilatés ou lorsque le nombre maximum d'itérations donné est atteint, l'algorithme est arrêté. Les résultats dépendent de la taille de l'élément structurant utilisé pour les opérations morphologiques.

Gillet applique des opérateurs de morphologie floue à l'histogramme couleur pour mettre en évidence les différents domaines modaux présents. Pour ce faire, il suppose que chaque point-couleur peut appartenir " plus ou moins " à un domaine modal. L'auteur définit le sousensemble flou caractérisé par sa fonction d'appartenance qui évalue le degré d'appartenance du point considéré à un domaine modal. Cette fonction est définie par une analyse de la concavité de l'histogramme couleur. Des transformations morphologiques floues sont ensuite appliquées à cette fonction d'appartenance afin de mettre en évidence les différents domaines modaux.

## 2.3 Segmentation et espace couleur

Le résultat d'une segmentation d'images couleur dépendant de l'espace couleur utilisé, de nombreux auteurs comparent les résultats de segmentation obtenu par la méthode qu'ils proposent en utilisant différents espaces couleur. Le tableau 2.2 montre les conclusions de quelques études réalisées sur les liens entre segmentation d'images couleur et espaces couleur. Nous ne présentons ici que des articles publiés dans des revues internationales mais d'autres études publiées dans des thèses ou des conférences peuvent compléter ce tableau 2.2. D'autre part, certains espaces cités dans ce tableau n'ont pas été présentés dans ce chapitre mais le lecteur pourra se référer à la publication citée pour plus de détails. Les espaces (U,V,W) et  $(U^*,V^*,W^*)$ sont ainsi des espaces perceptuellement uniformes qui sont notamment à l'origine des espaces  $(L^*, a^*, b^*)$  et  $(L^*, u^*, v^*)$  de la C.I.E.. L'espace (Y, Cb, Cr) est un espace de transmission des signaux de télévision numériques. L'espace (Y, u, v) est un système luminance-chrominance où l'axe de luminance correspond à l'axe achromatique du système (R, G, B) et le plan de chrominance correspond au plan défini par le triangle des couleurs du système (R, G, B). L'espace  $(Y, Q_{RG}, Q_{RB})$  est aussi un espace luminance-chrominance dans lequel  $Q_{RG}$  et  $Q_{RB}$  représentent respectivement les ratios de couleurs R/(R+G) et R/(R+B). Enfin, l'espace  $(y_1, y_2, y_3)$  est une normalisation de l'espace (I1, I2, I3) où les composantes I2 et I3 sont normalisées par I1 qui représente une composante de luminance dans le système d'Otha.

Les conclusions de ces études ne débouchent pas sur l'identification d'un espace qui est le mieux adapté à la segmentation des images. La conclusion est donc qu'il n'en existe pas. Les espaces couleur choisis dépendent de la méthode utilisée, de la nature de l'image à segmenter et des objectifs recherchés.

La recherche d'un espace couleur adapté est souvent réalisée a posteriori afin d'indiquer dans quel espace, la méthode proposée fournit les meilleurs résultats. En effet, il est très difficile a priori de déterminer quel est l'espace couleur qui permet la meilleure segmentation. Par ailleurs, les critères d'évaluation de la qualité des résultats fournis par un algorithme de segmentation varient d'une étude à l'autre.

### 2.4 Synthèse

La première partie de ce chapitre présente les espaces couleur les plus fréquemment utilisés en traitement d'images couleur. Ces derniers ont été présentés selon la classification proposée par Vandenbroucke qui les distinguent en quatre familles suivant leurs caractéristiques [Van00] :

- les espaces de primaires pour lesquelles la couleur est obtenue par un mélange de trois primaires,
- les espaces de luminance-chrominance composés d'une composante de luminance et de deux composantes de chrominance pour définir une couleur,
- les espaces perceptuels utilisant la teinte, la saturation et l'intensité pour définir la couleur,
- les espaces d'axes indépendants issus de méthodes statistiques afin de décorréler les axes.

Référence	Espaces comparés	Espace(s) choisi(s)	
[OPR78]	(R,V,B), (Y,I,Q), (I,S,T)	Т	
[OKS80]	$(R,V,B), (X,Y,Z), (Y,I,Q), (L^*,a^*,b^*),$	( <i>I</i> 1, <i>I</i> 2, <i>I</i> 3)	
	$(U^*, V^*, W^*), (I, r, g), (I, S, T), (I1, I2, I3)$		
[LL90]	(R,V,B), $(X,Y,Z),$ $(Y,I,Q),$ $(U,V,W),$	(I1, I2, I3), (R, V, B)	
	( <i>I</i> 1, <i>I</i> 2, <i>I</i> 3)		
[GH92]	$(R,V,B), (Y,I,Q), (L^*,a^*,b^*), (L,C,H)$	$(L^*, a^*, b^*), (L, C, H)$	
[AR96]	(R,V,B), (r,v,b), (I,S,T)	(I,S,T)	
[LR97]	$(R,V,B), (Y,u,v), (Y,Q_{RG},Q_{RB})$	(R,V,B)	
[BMB <sup>+</sup> 92]	$(R,V,B), (Y,I,Q), (L^*,a^*,b^*), (L^*,u^*,v^*),$	$(L^*, a^*, b^*), (L^*, u^*, v^*)$	
	(I,r,g), (I,S,T), (I1,I2,I3)		
[LY94]	$(R,V,B), (X,Y,Z), (Y,I,Q), (L^*,a^*,b^*),$	( <i>I</i> 1, <i>I</i> 2, <i>I</i> 3)	
	(I, S, T), (I1, I2, I3)		
[LCK94]	$(R,V,B), (X,Y,Z), (Y,I,Q), (L^*,a^*,b^*),$	$(y_1, y_2, y_3), (I, r, g)$	
	$(U^*, V^*, W^*), (I, r, g), (I, S, T), (I1, I2, I3),$		
	$(y_1, y_2, y_3)$		
[PR02]	$(R,V,B), (L^*,a^*,b^*), (L,C,H), (I,S,T),$	$(I1, I2, I3), (i_1, i_2, i_3)$	
	$(I1, I2, I3), (i_1, i_2, i_3)$		
[Muk02]	$(R,V,B), (Y,I,Q), (L^*,a^*,b^*), (U,V,W),$	( <i>I</i> 1, <i>I</i> 2, <i>I</i> 3)	
	( <i>I</i> 1, <i>I</i> 2, <i>I</i> 3)		
[DS05]	$(R,V,B), (r,v,b), (Y,Cb,Cr), (L^*,a^*,b^*),$	dépend de l'algorithme	
	(I,S,T), (I1,I2,I3)		
[PBC05]	$(R,V,B), (Y,Cb,Cr), (L^*,a^*,b^*), (I,S,T)$	Tous les espaces sont	
		jugés équivalent.	

Tableau 2.2 : Segmentation et espace couleur

Cette liste bien que non exhaustive met en avant la difficulté qu'il peut y avoir à utiliser et choisir un espace plutôt qu'un autre tant le nombre d'espaces couleur est important. Dans notre étude concernant la segmentation d'images couleur, nous utiliserons certains des espaces présentés afin de segmenter une image couleur dans un espace couleur adapté.

Aussi, après cette première partie traitant des principaux espaces couleur, nous avons présenté un état de l'art des méthodes de segmentation d'images couleur. Dans cette deuxième partie, nous avons préférentiellement présenté les méthodes de segmentation en régions et plus particulièrement les méthodes de segmentation analysant la distribution des points couleur dans un espace couleur puisque la méthode que nous proposons dans ce manuscrit appartient à cette famille.

Dans la dernière partie de ce chapitre, nous avons abordé l'influence du choix de l'espace de représentation sur la qualité des résultats obtenus par des méthodes de segmentation d'images couleur. De nombreux auteurs ont étudié les rapports entre les espaces de représentation de la couleur et l'analyse d'images couleur [TFMB04]. Beaucoup ont soulevé le problème du choix délicat de l'espace couleur pour la segmentation d'images. Notre synthèse montre que l'utilisation d'aucun espace couleur ne permet d'obtenir des résultats satisfaisants pour toute méthode de segmentation appliquée à tout type d'images.

Notre contribution va donc porter sur la comparaison des espaces couleur pour la segmentation d'images couleur. Nous proposons de nous placer dans un contexte non supervisé pour comparer les différents espaces de représentation de la couleur. La problématique que nous abordons consiste à déterminer ceux qui sont les mieux adaptés pour différencier les classes de pixels en présence dans l'image.

Notre contribution s'inscrivant dans un contexte de construction des classes de pixels par apprentissage non supervisé, nous nous sommes dirigés vers une segmentation d'images par multi-seuillage des composantes couleur d'un espace. La détection des seuils s'effectue par l'analyse des histogrammes mono-dimensionnels des trois composantes couleur associées à l'espace couleur.

Dans le chapitre 3, nous décrivons cette méthode de multi-seuillage en insistant sur les deux points clés de cette approche : le lissage de chacun des histogrammes mono-dimensionnels afin de mettre en évidence ses pics significatifs et la détection à proprement parler des seuils délimitant ses pics.

# **Chapitre 3**

## Segmentation par multi-seuillage

L'approche de segmentation proposée consiste en un multi-seuillage des composantes de l'image couleur, les seuils étant déterminés par l'analyse des histogrammes 1D des composantes couleur (cf. figure 3.1). Nous supposons que les points couleur des pixels composant les régions sont regroupés dans des "domaines modaux" de l'espace (R, V, B) à savoir des zones à forte densité de probabilité d'apparition. Le point clé de la segmentation consiste à trouver ces domaines modaux.

Par construction, les distributions des points couleur inclus dans ces domaines modaux sont projetées sur chacun des histogrammes 1*D* des trois composantes, donnant naissance à des *pics* dans ces histogrammes. Ces pics représentent les niveaux des pixels dont les tailles de population sont élevées tandis que les niveaux des pixels avec de faibles tailles de population correspondent à des *vallées*.

Nous effectuons le produit cartésien des intervalles délimités par les seuils détectés pour partitionner l'espace couleur 3D en parallélépipèdes rectangles. Les parallélépipèdes rectangles à forte densité de population de pixels seront les domaines modaux de l'espace couleur. A chaque domaine modal correspond une *classe* de pixels. Une fois les classes construites, il suffit alors d'effectuer une analyse en composantes connexes des pixels assignés aux mêmes classes pour obtenir une image segmentée en régions.

Le premier paragraphe traitera des histogrammes mono-dimensionnels des composantes couleur utilisés en vue de cette classification. Nous verrons alors que la détection des seuils par l'analyse directe de ces histogrammes ne fournit pas de résultats satisfaisants. Pour cette raison, nous détaillerons dans le deuxième paragraphe de ce chapitre l'opération de filtrage nécessaire à la détection des seuils. Le troisième paragraphe présentera comment les valeurs de seuils sont déterminées afin de délimiter les pics présents dans les histogrammes. Le quatrième paragraphe traitera de la construction des classes de pixels à partir des seuils détectés dans les trois histogrammes 1*D* de l'espace couleur (*R*, *V*, *B*) en vue de la segmentation de l'image.

Un point crucial de cette méthode est la qualité du filtrage effectué. Ainsi, le réglage des paramètres de lissage est une opération délicate. Le cinquième paragraphe détaillera différentes méthodes filtrant de manière adaptative les histogrammes pour mettre en évidence les pics. Enfin, nous appliquons notre approche à la segmentation d'images de référence. Ceci nous



Figure 3.1 : Multi-seuillage par analyse d'histogrammes 1D des composantes couleur.

permet de mettre en évidence les limites des performances atteintes par cette méthode quand l'espace de représentation des couleurs est uniquement l'espace (R, V, B).

## **3.1** Calcul d'histogrammes 1D couleur

Dans les approches de segmentation par classification pixellaire, les pixels de l'image sont projetés dans un espace couleur. Cette projection permet l'obtention d'un point couleur dont les coordonnées dans l'espace couleur sont celles des niveaux de couleur du pixel considéré. La figure 3.2 illustre la projection d'un pixel d'une image de synthèse dans l'espace couleur (R,V,B). La figure 3.2(a) montre le pixel P(93,101) de coordonnées (93,101) dans le plan image considéré pour illustrer une telle projection. Ce pixel dont les niveaux de rouge, vert et bleu valent respectivement  $C^R(P(93,101)) = 62$ ,  $C^V(P(93,101)) = 35$  et  $C^B(P(93,101)) = 30$  est représenté dans l'espace couleur (R,V,B) par le point C(P) de coordonnées (62,35,30). La figure 3.2(b) montre le point couleur C(P) dans l'espace couleur (R,V,B).

Généralement, les composantes couleur rouge, verte et bleue des images couleurs sont codées chacune sur 8 bits. Il y a donc 256 nuances de rouge, de vert et de bleu qui permettent de coder  $256^3 = 16777216$  couleurs possibles. Il est rare qu'une image utilise une telle palette de couleurs. Ainsi, les images couleur contiennent généralement un nombre restreint de couleurs comme le montre la figure 3.3. L'image 3.3(a) de cette figure est la même que l'image de synthèse de la figure 3.2(a) qui nous servira d'exemple pour illustrer ce chapitre. Cette image comprend 6 régions distinctes de couleurs différentes :



(a) Pixel P(93, 101) de niveau de composante  $C^{R}(P(93, 101)) = 62$ ,  $C^{V}(P(93, 101)) = 35$  et  $C^{B}(P(93, 101)) = 30$ .



(b) Projection du pixel P(93, 101) de l'image de synthèse en un point couleur C(P) dans l'espace couleur (R, V, B).



- le fond marron  $\Re 1$ ,
- un carré orange R2,
- une tache mauve  $\Re 3$ ,
- un carré jaune \$\$4,
- un anneau vert clair  $\Re 5$ ,
- un disque vert foncé  $\Re 6$ .



(a) Image de synhtèse.



(b) Points couleur correspondant aux pixels de l'image de synthèse dans l'espace (R,V,B).

**Figure 3.3 :** Histogramme 3D de l'image 3.3(a) dans l'espace (R, V, B).

La figure 3.3(b) montre l'histogramme 3*D* de l'image de synthèse 3.3(a) dans l'espace couleur (R, V, B). Sur cette figure, les densités de population des couleurs présentes dans l'image sont représentées par des sphères. Plus le diamètre de la sphère est important, plus le nombre d'occurences de pixels ayant cette couleur dans l'image est important. On peut ainsi remarquer sur cette figure 3.3(b) que les différentes régions de l'image de synthèse donnent naissance à des domaines modaux dans l'espace (R, V, B). De même, on peut remarquer qu'un grand nombre de couleurs dans l'espace (R, V, B) ne sont pas présentes dans l'image de la figure 3.3(a).

L'histogramme mono-dimensionnel associé à la composante couleur  $k \ (k \in \{R, V, B\})$  d'une image *I* recense la population des pixels caractérisés par les niveaux de la composante. Un histogramme mono-dimensionnel  $h^k[I]$  est une structure à une dimension à support borné composée de cellules  $h^k(C^k(P))$  avec  $0 \le C^k(P) \le 255$  tel que  $C^k(P) \rightarrow h^k(C^k(P))$  avec  $h^k(C^k(P)) =$  $card(C^k(P))$ .



Figure 3.4 : Histogrammes 1D de l'image de synthèse de la figure 3.3(a).

Le but de la segmentation consiste alors à détecter les seuils délimitant les pics afin de déterminer les domaines modaux de l'espace couleur. La figure 3.4 montre les histogrammes des composantes rouge (cf. figure 3.4(a)), verte (cf. figure 3.4(b)) et bleue (cf. figure 3.4(c)) de l'image de synthèse. A cause de la projection de la distribution des points couleur sur une des composantes, plusieurs domaines modaux peuvent former un seul pic de l'histogramme. Par exemple, l'histogramme de la composante bleue de l'image de synthèse ne possède que trois pics pour les six régions de l'image. Ainsi, le pic de plus grande amplitude correspond aux régions  $\Re 1$ ,  $\Re 2$ ,  $\Re 4$ , le pic de plus faible amplitude à la région  $\Re 5$  et le troisième pic aux régions  $\Re 3$  et  $\Re 6$  de l'image de synthèse.

Comme les couleurs des pixels sont perturbées par le bruit lié notamment aux imperfections de la chaîne d'acquisition, les histogrammes le sont également. Par conséquent, les histogrammes possèdent de nombreux *artefacts*. Il est donc primordial de filtrer les histogrammes afin de ne prendre en compte que les pics qui correspondent effectivement aux régions de l'image.

## **3.2** Filtrage d'histogrammes 1D par lissage Gaussien

Dans le cadre de la segmentation d'images, l'opération de filtrage d'histogrammes a pour objectif d'éviter une sur-segmentation ou une sous-segmentation en ne conservant que les *pics* 

*significatifs* qui correspondent effectivement aux régions de l'image à analyser. Il existe de nombreuses méthodes qui permettent le multi-seuillage par analyse d'histogrammes et qui intègrent des opérations de filtrage. Ces méthodes se divisent en six familles [SS04] :

- Les méthodes basées sur l'analyse des formes de l'histogrammes analysent les pics, les vallées et les courbures des histogrammes à partir d'histogrammes lissés.
- Les méthodes de clustering dans lesquelles l'histogramme est séparé en deux classes de pixels ou modélisé par un mélange de fonctions Gaussiennes.
- Les méthodes entropiques qui utilisent un critère d'entropie à partir de la distribution des niveaux de gris d'un histogramme.
- Les méthodes de mesure de similarité qui déterminent des seuils à partir de mesures de similarité de l'image originale et d'un ensemble d'images originales binarisées.
- Les méthodes spatiales qui utilisent la distribution des niveaux de gris en plus de l'information du voisinage des pixels.
- Les méthodes localement adaptées où les seuils sont calculés à partir des pixels et de caractéristiques locales de l'image.

Le filtrage d'histogramme a abouti à de nombreux travaux et il est délicat d'en faire une liste exhaustive ainsi qu'une classification. Dans ce chapitre, nous nous contenterons de présenter le principe de filtrage par lissage car c'est cette famille de méthodes que nous avons utilisée pour déterminer les pics significatifs des histogrammes.

Cette opération de lissage a pour objectif de fournir un histogramme lissé qui ne conserve que les pics significatifs de l'histogramme à analyser. Pour cela, les méthodes de lissage d'histogrammes consistent à appliquer une opération de convolution notée \* entre l'histogramme d'une composante couleur  $h^k(x)$  (avec  $k \in \{R, V, B\}$ ) et une fonction analysante g(x). Il existe plusieurs types de fonctions analysantes. Cependant, les fonctions analysantes utilisées en vue de lisser les histogrammes sont généralement des Gaussiennes car elles permettent de lisser plus ou moins fortement l'histogramme suivant l'écart-type  $\tau$  choisi. Cette fonction Gaussienne centrée en 0 et d'écart-type  $\tau$  aussi appelée *noyau Gaussien* sera notée  $g(x, \tau)$  telle que :

$$g(x,\tau) = \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{x^2}{2\tau^2}\right]$$
(3.1)

L'écart-type est un paramètre de ce noyau Gaussien qui permet de lisser plus ou moins l'histogramme, ce qui à pour conséquence de mettre en évidence un nombre plus ou moins grand de pics dans l'histogramme.

Par la suite, le résultat de la convolution entre l'histogramme d'une composante couleur  $h^k(x)$  et le noyau Gaussien  $g(x, \tau)$  noté  $h^k(x, \tau)$  s'exprime par :

$$h^{k}(x,\tau) = h^{k}(x) * g(x,\tau) = \int_{-\infty}^{+\infty} h^{k}(u) \cdot g(x-u,\tau) du = \int_{-\infty}^{+\infty} h^{k}(u) \cdot \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{(x-u)^{2}}{2\tau^{2}}\right] du.$$
(3.2)

La principale difficulté du lissage d'histogrammes tel que nous venons de le définir est le choix du paramètre  $\tau$ . Les figures 3.5, 3.6 et 3.7 illustrent l'influence de l'écart-type du noyau Gaussien sur le lissage de l'histogramme  $h^B(x)$  de la composante bleue de l'image de synthèse 3.3(a) respectivement pour des valeurs  $\tau = 2$ ,  $\tau = 5$  et  $\tau = 10$ .



**Figure 3.5 :** Convolution de l'histogramme de la composante bleue  $h^{\mathcal{B}}(x)$  de l'image 3.3(a) avec un noyau gaussien d'écart- type  $\tau = 2$ .

La figure 3.5(c) montre le résultat de la convolution entre l'histogramme  $h^B(x)$  et un noyau Gaussien d'écart-type  $\tau = 2$  représenté sur la figure 3.5(b). Dans ce cas, le lissage a correctement été effectué puisque d'une part les pics significatifs de l'histogramme sont conservés et que d'autre part les artefacts provenant du bruit de l'image sont éliminés.



**Figure 3.6 :** Convolution de l'histogramme de la composante bleue  $h^{\mathcal{B}}(x)$  de l'image 3.3(a) avec un noyau gaussien d'écart-type  $\tau = 5$ .

La figure 3.6(c) montre le résultat de la convolution entre l'histogramme  $h^B(x)$  et un noyau Gaussien d'écart-type  $\tau = 5$  représenté sur la figure 3.6(b). Nous pouvons remarquer sur la figure 3.6(c) que le pic correspondant à la région  $\Re 5$  et celui correspondant aux régions  $\Re 3$ et  $\Re 6$  ont fusionné avec l'accroissement de l'écart-type. L'histogramme lissé correspond à un *sur-lissage* car des pics significatifs ont été fusionnés.



**Figure 3.7 :** Convolution de l'histogramme de la composante bleue  $h^{\beta}(x)$  de l'image 3.3(a) avec un noyau gaussien d'écart-type  $\tau = 10$ .

La figure 3.7(c) montre le résultat de la convolution entre l'histogramme  $h^B(x)$  et un noyau Gaussien d'écart-type  $\tau = 10$  représenté sur la figure 3.7(b). Cette figure illustre également le cas d'un histogramme sur-lissé pour les mêmes raisons que le cas précédent. En outre, on peut y remarquer que les frontières délimitant les pics sont différentes de celles de l'histogramme original. Cette figure met ainsi en évidence que le résultat d'un produit de convolution peut modifier la localisation sur l'axe des abscisses des *seuils* délimitant les pics des histogrammes. Ce phénomène est d'autant plus important que l'écart-type du noyau Gaussien est grand.

Une fois le lissage des histogrammes effectué, il est nécessaire d'identifier les différents pics présents dans ces histogrammes en détectant les seuils qui les délimitent.

## **3.3 Détection des seuils**

Dans la littérature, différentes méthodes permettent de déterminer des seuils pour délimiter les pics en vue de la segmentation d'images par multi-seuillage. Le premier paragraphe a pour objectif de définir ces seuils tandis que les deux paragraphes suivants ont pour objectifs de montrer l'influence de la détection des seuils en vue de segmenter une image en niveaux de gris ou une image couleur.

#### **3.3.1** Définition des seuils

Une fois les histogrammes lissés, il convient de déterminer les seuils délimitant les pics d'un histogramme représenté par une courbe généralement constituée de zones convexes ou concaves. Il est possible de caractériser les pics de l'histogramme par la détection des points d'inflexion ou des extrema locaux de cette courbe. Les zones où la courbe est concave correspondent aux pics tandis que les zones où la courbe est convexe correspondent aux vallées. La figure 3.8 montre un histogramme constitué de deux Gaussiennes centrées respectivement en  $\mu_1 = 64$  et  $\mu_2 = 192$  et d'écart-type  $\tau = 30$ . Cet histogramme comprend donc deux pics numérotés 1 et 2 sur la figure 3.8. Les valeurs d'abscisse des amplitudes maximales correspondant au plus grand nombre d'occurrences au sein d'un pic sont notées  $A_i^k$  pour le  $i^{\text{ème}}$  pic de la



Figure 3.8 : Caractéristiques des histogrammes.

composante k. Les valeurs de seuils qui permettent de caractériser les pics seront notées  $S_{g,i}^k$  et  $S_{d,i}^k$  lorsqu'ils sont situés respectivement à gauche ou à droite du  $i^{\text{ème}}$  pic de la composante k. Différentes méthodes existent pour déterminer des seuils délimitant les pics.

En fonction de la méthode choisie, les seuils  $S_{g,i}^k$  et  $S_{d,i}^k$  peuvent alors être égales soit aux points d'inflexion soit aux minima locaux. Ainsi, le premier paragraphe explique comment sont déterminés les minima locaux des histogrammes tandis que le second paragraphe explique comment sont déterminés les points d'inflexion des histogrammes.

#### 3.3.1.1 Détection des minima locaux

Les valeurs d'abscisse des minima locaux servant à délimiter les pics des histogrammes lissés  $h^k(x, \tau)$  peuvent être déterminés à partir de l'analyse des passages par zéro de l'estimation de la dérivée première  $h_x^k(x, \tau) = \frac{\partial}{\partial x} h^k(x, \tau)$ . Les minima locaux délimitant les pics sont notés  $T_{g,i}^k$  lorsqu'ils sont situés à gauche de l'amplitude maximale du  $i^{\text{ème}}$  pic de la composante k et  $T_{d,i}^k$ lorsqu'ils sont situés à droite. La figure 3.9 montre la représentation graphique de l'estimation de la dérivée première  $h_x^k(x, \tau)$  de l'histogramme  $h^k(x)$  de la figure 3.8.



**Figure 3.9 :** Dérivée première  $h_x^k(x, \tau)$  de l'histogramme  $h^k(x, \tau)$  de la figure 3.8.

On peut remarquer sur cette figure que :

- Les maxima locaux correspondent aux passages par zéro de l'estimation de la dérivée première dont le signe passe du signe + au signe -.
- Les minima locaux correspondent aux passages par zéro de l'estimation de la dérivée première dont le signe passe du signe – au signe +.

Seule les extrémités de l'histogramme sont traitées différemment pour déterminer les minima locaux d'un histogramme  $h_x^k(x, \tau)$ . Dans l'exemple de la figure 3.8, on peut également remarquer que pour le 1<sup>er</sup> pic et pour son pic voisin de droite (le 2<sup>ème</sup> pic), le seuil droit du 1<sup>er</sup> pic et le seuil gauche du 2<sup>ème</sup> pics sont confondus. On a par conséquent l'égalité  $T_{d,1}^k = T_{g,2}^k$ . De manière plus générale, on a l'égalité  $T_{d,i}^k = T_{g,i+1}^k$ .

#### 3.3.1.2 Détection des points d'inflexion

Les valeurs d'abscisse des points d'inflexion des histogrammes lissés  $h^k(x,\tau)$  peuvent être déterminés à partir des passages par zéro de l'estimation de la dérivé seconde  $h_{xx}^k(x,\tau) = \frac{\partial^2}{\partial x^2}h^k(x,\tau)$ . Les points d'inflexion délimitant les pics sont notés  $I_{g,i}^k$  lorsqu'ils sont situés à gauche de l'amplitude maximale du  $i^{\text{ème}}$  pic de la composante ket  $I_{d,i}^k$  lorsqu'ils sont situés à droite.



La figure 3.10 montre la représentation graphique de l'estimation de la dérivé seconde  $h_{xx}^k(x,\tau)$  de l'histogramme  $h^k(x,\tau)$  de la figure 3.8.

**Figure 3.10 :** Dérivée seconde  $h_{xx}^k(x, \tau)$  de l'histogramme  $h^k(x, \tau)$  de la figure 3.8.

On peut remarquer sur cette figure que :

- Les points d'inflexion situés à gauche de l'amplitude maximale d'un pic correspondent au passage par zéro de la dérivé seconde dont le signe passe du signe + au signe -.
- Les points d'inflexion situés à droite de l'amplitude maximale d'un pic correspondent au passage par zéro de la dérivé seconde dont le signe passe du signe – au signe +.

Afin d'expliquer l'influence des seuils utilisés pour délimiter les pics pour segmenter une image, nous utiliserons une nouvelle image de synthèse. Cette image représentée sur la figure 3.11(a) a été construite à partir de deux rectangles. Les niveaux des composantes couleur R, V et B du rectangle gris foncé de gauche ont été initialement fixés à 64 et les niveaux des composantes couleur du rectangle gris clair de droite ont été initialement fixés à 192. Pour obtenir l'image de synthèse de la figure 3.11(a), un bruit blanc Gaussien a été ajouté sur les trois composantes couleur des pixels de l'image. Les figures 3.11(b), 3.11(c) et 3.11(d) montrent respectivement les histogrammes des composantes rouge, verte et bleue de cette image. L'image de la figure 3.11(a) a ainsi été construite de telle sorte que ces histogrammes puissent être modélisés à partir de deux Gaussiennes  $g_1(x, 30)$  et  $g_2(x, 30)$  centrées respectivement en  $\mu_1 = 64$ 

et  $\mu_2 = 192$  et d'écart-type  $\tau = 30$ . Sur les figures 3.11(b), 3.11(c) et 3.11(d), la première Gaussienne  $g_1(x, 30)$  a été superposée aux histogrammes en rouge tandis que la seconde  $g_2(x, 30)$  a été superposée en vert. De même, les valeurs des points d'inflexion et des minima locaux apparaissent sur la figure 3.11.



Figure 3.11 : Image de synthèse avec les histogrammes des composantes R, V et B et une modélisation des histogrammes à partir des deux gaussiennes  $g_1(x, 30)$  et  $g_2(x, 30)$ .

Afin d'illustrer l'influence des seuils pour construire les classes de pixels, la segmentation par classification de pixels de l'image de la figure 3.11(a) a été effectuée pour l'image couleur et pour l'image de la composante rouge suivant deux cas :

- 1<sup>er</sup> cas : les seuils caractérisant les pics correspondent aux points d'inflexion,
- 2<sup>ème</sup> cas : les seuils caractérisant les pics correspondent aux minima locaux.

# **3.3.2** Influence du choix des seuils pour la segmentation d'une image en niveaux de gris

La figure 3.12 illustre l'influence des seuils pour une image en niveaux de gris suivant que les seuils utilisés sont les points d'inflexion ou les minima locaux. A titre d'exemple, cette image à niveaux de gris est l'image de la composante rouge de l'image de synthèse de la figure 3.11(a). Les pixels étiquetés en rouge de l'image 3.12(a) correspondent aux pixels dont les niveaux de rouge sont compris entre les points d'inflexion  $I_{g,1}^R = 34$  et  $I_{d,1}^R = 94$ . Les pixels étiquetés en vert correspondent à ceux dont les niveaux de rouge sont compris entre  $I_{g,2}^R = 162$  et  $I_{d,2}^R = 192$ . Enfin, les pixels étiquetés en noir correspondent aux pixels de l'image 3.12(a) dont les niveaux de rouge n'appartiennent à aucune de ces deux pixels.



(a) Pixels de l'image 3.11(a) dont les niveaux de gris appartiennent aux domaines modaux de l'histogramme de la figure 3.11(b) modélisé par deux gaussiennes.



(b) Pixels de l'image 3.11(a) dont les niveaux de gris appartiennent aux domaines modaux limités par les creux des vallées de l'histogramme de la figure 3.11(b) modélisé par deux gaussiennes.

Figure 3.12 : Influence du choix des seuils pour une image en niveaux de gris.

Pour l'image 3.12(b), les pixels ont été assignés à une classe de pixels à partir de seuils correspondant aux minima locaux. Les pixels étiquetés en rouge de l'image 3.12(b) correspondent aux pixels dont les niveaux de rouge sont compris entre les minima locaux du pic modélisé par la Gaussienne  $g_1(x, 30)$  c'est-à-dire les pixels dont le niveau de rouge est compris entre  $T_{g,1}^R = 0$ et  $T_{d,1}^R = 128$ . Les pixels en vert correspondent à ceux dont le niveau de rouge est compris entre les minima locaux du pic modélisé par la Gaussienne  $g_2(x, 30)$  (entre  $T_{g,2}^R = 128$  et  $T_{d,2}^R = 255$ ).

On peut remarquer sur l'image 3.12(a) que tous les pixels étiquetés à la classe rouge ou la classe verte ont été correctement assignés contrairement aux pixels de l'image 3.12(b). La figure 3.13 explique cette erreur de classification. La figure 3.13(a) montre l'histogramme de la composante rouge de l'image 3.11(a) sur laquelle ont été superposées les Gaussiennes  $g_1(x, 30)$  et  $g_2(x, 30)$ . La valeur de seuil 128 qui correspond au minimum local délimitant les 2 pics de l'histogramme correspond également à la valeur d'abscisse à laquelle se croise les deux

Gaussiennes. Si les seuils utilisés sont les minima locaux alors les pixels dont les niveaux de rouge sont compris dans la zone remplie en rouge sur la figure 3.13(a) de la Gaussienne  $g_1(x, 30)$ seront mal classés dans l'image puisqu'ils sont situés à une valeur d'abscisse supérieure au minimum local délimitant les deux pics. Pour la même raison, les pixels dont les niveaux de rouge sont compris dans la zone remplie en vert sur la figure 3.13(a) de la Gaussienne  $g_2(x, 30)$ seront mal classés dans l'image. Ces zones qui symbolisent l'erreur de classification signifient que plus leurs surfaces sont élevées, plus le nombre de pixels mal classés sera important. Aussi, en prenant comme seuils les points d'inflexion, la figure 3.13(b) montre que ces surfaces sont beaucoup plus faibles et justifie par conséquent le taux de bonne classification des pixels lorsque les points d'inflexion sont utilisés comme seuils.



(a) Erreur de classification si les valeurs seuils uti- (b) Erreur de classification si les valeurs seuils utilisés sont les creux des vallées.

lisés sont les points d'inflexion.

Figure 3.13 : Erreur de classification suivant les valeurs de seuils utilisés.

Si le taux de bonne classification est meilleur en détectant les points d'inflexion, il faut remarquer que tous les pixels ne sont pas étiquetés contrairement au cas de la détection des minima locaux. Les algorithmes de segmentation exploitant les points d'inflexion utilisent alors une approche dite *coarse to fine*. Ce type d'approche se déroule en deux étapes. La première étape consiste en une segmentation grossière (coarse) en étiquetant les pixels à partir des points d'inflexion détectés. Il succède alors une segmentation fine (fine) afin d'étiqueter les pixels qui n'appartiennent à aucune classe issue de la segmentation grossière. La première segmentation permet de réduire significativement la quantité de données à traiter pour la seconde segmentation et par conséquent de réduire le temps de traitement de cette 2<sup>ème</sup> étape.

Après avoir analysé l'influence du critère de détection des seuils pour la segmentation d'une image en niveaux de gris, le paragraphe suivant tentera d'expliquer l'apport du caractère tridimensionnel de la couleur pour détecter les seuils en vue de la segmentation d'une image couleur.

#### 3.3.3 Influence du choix des seuils pour la segmentation d'une image couleur

En effet, la prise en compte du caractère tri-dimensionnel de la couleur permet de limiter une mauvaise classification puisque la construction d'une classe sera effectuée à partir de seuils provenant des histogrammes 1D de trois composantes. Nous utilisons les trois composantes couleur de l'espace couleur (R, V, B) en vue d'une telle classification (c.f. figure 3.14).

Les pixels étiquetés en rouge de l'image 3.14(a) correspondent aux pixels de l'image 3.11(a) dont les points couleur appartiennent aux domaines modaux de l'espace 3D déterminés par le produit cartésien des intervalles  $\begin{bmatrix} I_{g,1}^R, I_{d,1}^R \end{bmatrix}$ ,  $\begin{bmatrix} I_{g,1}^V, I_{d,1}^V \end{bmatrix}$  et  $\begin{bmatrix} I_{g,1}^B, I_{d,1}^B \end{bmatrix}$ . Les seuils utilisés dans ce cas pour effectuer le produit cartésien sont les points d'inflexion des histogrammes 1D. De même, les pixels étiquetés en vert de l'image 3.14(a) correspondent aux pixels de l'image 3.11(a) dont les points couleur R, V et B appartiennent aux domaines de l'espace 3D déterminés par le produit cartésien des intervalles  $\begin{bmatrix} I_{g,2}^R, I_{d,2}^R \end{bmatrix}, \begin{bmatrix} I_{g,2}^V, I_{d,2}^V \end{bmatrix}$  et  $\begin{bmatrix} I_{g,2}^B, I_{d,2}^B \end{bmatrix}$ .

Pour l'image 3.14(b), les pixels ont été assignés à une classe de pixels suivant des valeurs de seuils correspondant aux minima locaux des histogrammes 1D. Les pixels étiquetés en rouge dans l'image 3.14(b) correspondent aux pixels dont les niveaux de rouge, de vert et de bleu sont compris entre les minima locaux du pic modélisé par la Gaussienne  $g_1(x, 30)$ , c'est-à-dire les pixels dont les niveaux des composantes couleur sont compris entre  $T_{g,1}^R = T_{g,1}^V = T_{g,1}^B = 0$  et  $T_{d,1}^R = T_{d,1}^V = T_{d,1}^B = 128$ . Les pixels étiquetés en vert correspondent à ceux dont les niveaux des composantes couleur sont compris entre les minima locaux du pic modélisé par la Gaussienne  $g_2(x,30)$  (entre  $T_{g,2}^R = T_{g,2}^V = T_{g,2}^B = 128$  et  $T_{d,2}^R = T_{d,2}^V = T_{d,2}^B = 255$ ).



leur appartiennent aux domaines modaux de l'es- leur appartiennent aux domaines modaux de l'espace (R, V, B) délimités par les points d'inflexion pace (R, V, B) délimités par les minima locaux dédétectés.

(a) Pixels de l'image 3.11(a) dont les points cou- (b) Pixels de l'image 3.11(a) dont les points coutectés.

**Figure 3.14 :** Influence du choix des seuils pour la segmentation de l'image couleur de la figure 3.11(a).

On peut remarquer pour ces deux images que tous les pixels classés ont été affectés correctement et que le nombre de pixels bien classés dans l'image 3.14(b) est supérieur à celui de l'image 3.14(a). On peut aussi remarquer pour une image couleur que la détection des minima locaux ne garantit pas la classification de tous les pixels contrairement au cas d'une image en niveaux de gris. Pour expliquer ce phénomène, la figure 3.15 montre les points couleur des pixels de l'image 3.11(a) projetés sur le plan couleur (R, V). La détection des minima locaux permet de partitionner ce plan couleur en 4 cadrans numérotés sur la figure 3.15 de 1 à 4 et délimités par des lignes bleues correspondent aux seuils. Aussi, dans ce plan couleur (R, V) seuls les pixels dont les points couleur associés sont situés dans les cadrans 1 et 3 seront étiquetés car seuls ces cadrans correspondent aux domaines modaux.



**Figure 3.15 :** Partitionnement du plan couleur (R, V).

En conclusion, le choix des critères à utiliser pour détecter les seuils est un débat ouvert. Mais, ces exemples semblent suggérer que l'utilisation des points d'inflexion semble être un choix plus judicieux pour la segmentation d'images en niveaux de gris car ils permettent de réduire le risque de mauvaise classification. Par ailleurs, l'utilisation des minima locaux semble être un choix judicieux pour la segmentation d'images couleur car leur caractère multidimensionnel semble apporter un meilleur compromis entre nombre de pixels classés et taux de bonne classification.

Les minima locaux des histogrammes 1D sont alors utilisés par la suite pour délimiter les domaines modaux de l'espace (R, V, B). A chaque domaine modal correspond une classe de pixels.

## 3.4 Construction d'une classe

La segmentation d'images couleur par analyse d'histogrammes mono-dimensionnels couleur consiste à déterminer des classes de pixels à partir de seuils délimitant les pics. Les seuils délimitant les pics permettent de partitionner l'espace couleur (R, V, B) en domaines modaux ayant la forme de parallélépipèdes rectangles.

La partition de l'espace couleur (R, V, B) permet ainsi d'obtenir  $N_R \times N_V \times N_B$  classes de pixels possibles où  $N_R$ ,  $N_V$  et  $N_B$  sont respectivement les nombres de pics significatifs déterminés à partir des histogrammes filtrés des composantes couleur R, V et B. Les seuils délimitant les pics significatifs sont notés  $S_{g,i}^k$  pour le seuil gauche et  $S_{d,i}^k$  pour le seuil droit du  $i^{\text{ème}}$  pic de la composante k. Les pixels de l'image appartenant à une classe ainsi construite vérifient le système d'équations 3.3 :

$$\begin{cases} S_{g,i}^{R} < C^{R}(P) < S_{d,i}^{R} \\ \text{et} \\ S_{g,l}^{V} < C^{V}(P) < S_{d,l}^{V} \\ \text{et} \\ S_{g,m}^{B} < C^{B}(P) < S_{d,m}^{B} \end{cases}$$
(3.3)

où *i*, *l* et *m* correspondent respectivement à un pic significatif des histogrammes  $h^R(x, \tau)$ ,  $h^V(x, \tau)$  et  $h^B(x, \tau)$  et sont limités par  $1 \le i \le N_R$ ,  $1 \le l \le N_V$  et  $1 \le m \le N_B$ . Pour une segmentation par classification pixellaire, il est impératif de ne sélectionner pour les  $N_R \times N_V \times N_B$  domaines possibles que les domaines modaux les plus représentatifs. Dans notre approche, ces domaines correspondent aux parallélépipèdes rectangles dans lesquelles les tailles de population de pixels les constituant sont les plus grandes.

La figure 3.16 illustre la construction de la classe de pixels de la région  $\Re 1$  correspondant au fond de l'image de synthèse 3.3(a) dans l'espace couleur (R, V, B). Cette classe de pixels est représentée par un parallélépipède rectangle délimité en bleu dans l'espace couleur (R, V, B) sur la figure 3.16(a). Les pixels dont les points couleur associés sont inclus dans ce parallélépipède rectangle appartiennent ainsi à la classe du fond. Ces pixels ont été laissés avec leurs points couleur d'origine dans l'image 3.16(b) et correspondent ainsi aux pixels appartenant à la classe construite de l'image. Ceux n'y appartenant pas ont été étiquetés en noir.

Les parallélépipèdes rectangles ainsi sélectionnés permettent de construire les  $N_C$  classes de pixels de l'image à segmenter. Les  $N_C$  classes construites permettent de classer les pixels de l'image en affectant  $N_C$  étiquettes différentes aux pixels P de l'image I à segmenter en fonction de leurs composantes couleur R, V et B.

Une fois les pixels assignés aux classes construites, une analyse en composantes connexes des pixels dans le plan image permet d'obtenir une partition de l'image en régions. En effet, une même classe de pixels peut correspondre à plusieurs régions différentes dans l'image avec des couleurs similaires. L'analyse en composantes connexes permet alors de distinguer ces régions et d'obtenir ainsi une segmentation telle [Zuc76].

Après avoir expliqué le principe de la segmentation d'images couleur par analyse d'histo-



(a) Représentation de la classe construite dans l'espace (R, V, B) de la région  $\Re 1$ .



(b) Pixels de l'image dont les niveaux de couleur appartiennent à la classe représentée sur la figure 3.16(a).

**Figure 3.16 :** Représentation de la classe construite dans l'espace (R, V, B) de la région  $\Re 1$  ainsi que des pixels de l'image dont les niveaux de couleur sont inclus dans le parallélépipède associé à cette classe.

grammes 1*D* couleur, nous allons nous intéresser plus précisément au filtrage des histogrammes par des méthodes de lissage. L'opération de filtrage est une étape essentielle pour ce type de méthodes puisqu'elles reposent sur la détection de pics significatifs.

## **3.5** Lissage multi-échelles d'histogrammes 1D

Nous présentons dans les paragraphes suivants deux méthodes de lissage d'histogrammes utilisant un produit de convolution entre l'histogramme d'une composante couleur de l'image et un noyau Gaussien. La première méthode, proposée par Lin et al. est une méthode heuristique permettant de lisser les histogrammes rapidement par exploration de l'histogramme [LWY96]. La deuxième méthode est le *Scale-Space Filtering* introduite par Witkin en 1984. Elle permet de déterminer précisément les valeurs des seuils délimitant les pics d'un histogramme [Wit84]. Ainsi, les deux premiers paragraphes de cette partie présenteront respectivement la méthode heuristique et le scale-space filtering. Nous mettons alors en évidence les limites atteintes par cette méthode. Nous proposerons alors un nouveau critère pour sélectionner les pics significatifs des histogrammes dont l'originalité consiste à prendre en compte l'arrangement spatial des niveaux dans l'image. Le dernier paragraphe présentera alors quelques résultats obtenus avec le critère proposé.

#### 3.5.1 Lissage Gaussien heuristique

Une méthode heuristique a été proposée afin de déterminer automatiquement l'écart-type du noyau Gaussien utilisé pour lisser l'histogramme dans l'objectif de ne conserver que les pics significatifs [LWY96]. Cette méthode est basée sur un algorithme itératif qui permet en un nombre restreint d'itérations (3 ou 4 généralement), de fournir un histogramme lissé. Cet algorithme consiste à lisser l'histogramme par une fonction Gaussienne à l'aide d'un critère basé sur le nombre de pics détectés. Tant que ce critère n'est pas vérifié, l'écart-type de la Gaussienne est augmenté selon une règle de mise à jour et l'histogramme est à nouveau lissé pour ne conserver que les pics significatifs.

#### 3.5.1.1 Principe

Cet algorithme de lissage consiste à déterminer l'écart-type du noyau Gaussien pour lisser un histogramme  $h^k(x)$  par itérations successives. Pour chaque itération, l'histogramme lissé  $h^k(x, \tau_i)$  est calculé avec la valeur de l'écart-type  $\tau_i$  où *i* est le numéro de l'itération. A la première itération i = 0, la valeur de l'écart-type vaut  $\tau_0 = 1$ . Pour les itérations suivantes, la valeur de l'écart-type est mise à jour par la relation  $\tau_{i+1} = \sqrt{d_{i+1}}$  ou  $d_{i+1}$  est la plus grande distance euclidienne entre deux pics voisins. Cette distance  $d_{i+1}$  prise entre les valeurs d'abscisses pour lesquelles l'amplitude des deux pics voisins est maximale, permet de caractériser la distribution des pics dans l'histogramme. En effet, si  $d_{i+1}$  est grande par rapport aux autres distances séparant deux pics voisins, on suppose alors que convoluer l'histogramme avec un noyau Gaussien d'écart-type plus grand permettra de supprimer les pics voisins qui sont proches les uns des autres correspondant à des artefacts et par conséquent de ne conserver que les pics significatifs de l'histogramme. Ainsi, au fur et à mesure des itérations, le nombre de pics présents dans l'histogramme à analyser diminue, ce qui a pour conséquence d'augmenter la distance entre deux pics voisins. La fonction liant l'écart-type de la Gaussienne et la distance maximale entre deux pics étant strictement monotone croissante, elle implique que la valeur de l'écart-type croit au fur et à mesure des itérations.

Un critère basé sur le nombre de pics détecté par la détermination des minima locaux permet alors d'arrêter l'algorithme. Trois cas déterminent si les pics détectés sont des pics significatifs :

- L'histogramme lissé possède un seul pic (histogramme mono-modal) ou est uniforme (pas de pic). Il est alors inutile de lisser avec une Gaussienne dont la valeur de l'écart-type est plus grande que la valeur courante puisqu'on obtiendra de nouveau un seul pic.
- Le nombre de pics à l'itération suivante est supérieur au nombre de pics détectés à l'itération courante. Dans ce cas, lisser avec un écart-type plus grand que l'écart-type courant fera apparaître de nouveaux pics au lieu d'en diminuer le nombre.
- Le nombre de pics à l'itération suivante est égal au nombre de pics détectés à l'itération courante. Dans ce cas, lisser avec un écart-type plus grand ne supprimera aucun autre pic. Les pics détectés montrent ainsi une certaine robustesse au lissage puisque leur nombre reste constant. Les pics présents dans l'histogramme sont alors considérés comme étant les pics significatifs de l'histogramme.

Les auteurs de cette méthode proposent deux variantes de l'algorithme. La première tente de déterminer un écart-type permettant de lisser l'histogramme dans son intégralité. Pour cette raison, nous l'appelons méthode globale. La seconde approche que nous appelons méthode locale permet de déterminer un ensemble d'écart-types adaptés localement à la dynamique de l'histogramme. Les deux paragraphes suivants présentent ces deux méthodes.

#### 3.5.1.2 Lissage global

Le schéma de la figure 3.17 présente cette méthode.



Figure 3.17 : Méthode de lissage global. Chaque étape est numérotée de 1 à 6.

A l'étape 1, l'écart-type  $\tau$  est initialisé à 1. Il correspond ainsi au pas de discrétisation de l'axe des abscisses. Cet écart-type sera mis à jour de manière heuristique à l'étape 5 par la formule :  $\tau_{i+1} = \sqrt{d_{i+1}}$ .

A la première itération i = 0 de l'algorithme de la figure 3.17, deux cas peuvent se produire. Si le nombre de pics  $p_1$  détectés de l'histogramme lissé est inférieur ou égal à 1, l'algorithme s'arrête parce que l'histogramme est respectivement uniforme  $(p_1 = 0)$  ou mono-modal  $(p_1 = 1)$ . Sinon, la valeur de l'écart-type  $\tau_0$  est mise à jour à la valeur  $\tau_1$  à l'étape 5 car  $p_1 \le (p_0 = \infty)$ . Le numéro de l'itération est incrémenté et un nouveau lissage est effectué à l'étape 2.

A la deuxième itération, la distance entre deux pics voisins vaut au moins 2 à cause du pas de discrétisation de l'axe des abcisse qui vaut 1. Ceci implique que la valeur  $\tau_1 = (\sqrt{d_1} \ge \sqrt{2})$  sera toujours plus grande que  $\tau_0 = 1$  et l'algorithme retourne à l'étape 2. Généralement, le nombre de pics détectés à la deuxième itération est inférieur au nombre de pics détectés lors de la première itération. L'algorithme retourne alors à l'étape 2 après avoir calculé  $\tau_1$  et incrémenté *i*.

A la  $i^{\text{ème}}$  itération  $i \ge 2$ , le nombre de pics  $p_{i+1}$  détectés est calculé à l'étape 3. Trois cas peuvent alors se produire :

- $p_{i+1} < p_i$ , l'algorithme continue à l'étape 4 pour une nouvelle itération puisque l'histogramme est de plus en plus lissé.
- $p_{i+1} = p_i$ , le nombre de pics est constant entre les itérations *i* et *i*+1. Les pics ainsi déterminés sont présents pour plusieurs valeurs d'écart-type consécutives et sont considérés comme étant des pics significatifs de l'histogramme. L'algorithme va alors à l'étape 5 pour calculer l'histogramme lissé avec le noyau Gaussien  $g(x, \tau_i)$ .
- $-p_{i+1} > p_i$ , le nombre pics détectés à l'itération *i* est plus grand que le nombre de pics détectés à l'itération *i* + 1. De nouveaux pics apparaissent et la valeur de l'écart-type du noyau Gaussien ne peut pas augmenter entre ces deux itérations. L'algorithme va alors à l'étape 5 pour calculer l'histogramme lissé avec le noyau Gaussien  $g(x, \tau_i)$ .

L'inconvénient majeur de cet algorithme est qu'il n'utilise qu'un seul écart-type pour lisser l'histogramme. Or la dynamique d'un histogramme peut varier selon les niveaux des composantes. Ceci peut avoir pour conséquence de fusionner certains pics ou de laisser des artefacts. Pour cette raison Lin propose une version locale de l'algorithme permettant de lisser l'histogramme  $h^k(x)$  avec des Gaussiennes d'écart-types adaptées localement à la dynamique de l'histogramme [LWY96].

#### 3.5.1.3 Lissage local

Le deuxième algorithme permettant de lisser localement l'histogramme est présenté sur la figure 3.18 :



Figure 3.18 : Méthode de lissage local. Chaque étape est numérotée de 1 à 6.

La première étape consiste à appliquer le premier algorithme de la figure 3.17 de manière à obtenir l'histogramme lissé  $h_1^k(x, \tau)$ . Ensuite, dans l'étape 2 on calcule à partir de l'histogramme lissé  $h_1^k(x, \tau)$  le nombre r de fois que cet histogramme lissé passe par l'amplitude moyenne défini comme étant la moitié de l'amplitude maximale, afin de diviser en r + 1 intervalles l'histogramme  $h^k(x)$  à partir des r intersections ainsi déterminées. L'étape 4 consiste à créer r + 1 histogrammes à partir de l'histogramme initial  $h^k(x)$  qui correspondent aux r + 1 intervalles définis par les r intersections. L'algorithme de la figure 3.17 est alors appliqué sur ces r + 1

histogrammes en étape 5 permettant ainsi de prendre en compte la dynamique locale de l'histogramme dans chacun de ces intervalles. Les r + 1 histogrammes lissés sont alors ré-assemblés pour obtenir l'histogramme lissé final  $h_2^k(x, \tau)$  en étape 6.

L'inconvénient majeur des méthodes heuristiques est que les résultats ne sont pas toujours satisfaisants. Ainsi, cette méthode a parfois tendance à lisser de manière trop importante les histogrammes. Ce sur-lissage peut parfois décaler la localisation des seuils délimitant les pics d'un histogramme et même fusionner des pics significatifs de l'histogramme. Contrairement à cette méthode heuristique, la méthode du scale-space filtering que nous allons détailler dans le paragraphe suivant permet de déterminer de manière plus précise les seuils délimitant les pics d'un histogramme. Il ne reste plus alors qu'à déterminer les pics significatifs à partir d'un critère spécifique.

#### 3.5.2 Scale-space filtering

Comme le souligne Lindeberg, les objets observés ne sont reconnaissables que pour une certaine plage d'échelles [Lin90, Lin94]. Par exemple, on ne peut distinguer visuellement les branches d'un arbre que si l'on ne se situe qu'à quelques mètres de celui-ci. La distance séparant l'objet de l'observateur joue alors le rôle du paramètre d'échelle. En appliquant ce concept à l'analyse d'histogrammes, on considère que les pics significatifs de l'histogramme ne sont détectables que pour une certaine plage d'échelles. Quand on examine l'histogramme de trop près, on peut distinguer les artefects tandis qu'en l'observant avec une échelle importante on ne distinguera que les pics significatifs.

Cette démarche est suivie par le scale-space filtering (SSF) afin de détecter les seuils délimitant les pics significatifs des histogrammes. Pour cela, l'histogramme à analyser est lissé avec des noyaux Gaussiens d'écart-types variables  $g(x, \tau)$  afin d'en déterminer des seuils. On parle ainsi de *grandes échelles* (ou hautes échelles) pour des valeurs d'écart-type importantes et de *petites échelles* (ou basses échelles) pour des valeurs d'écart-type faibles.

L'idée originale de Witkin [Wit84] consiste à lisser l'histogramme avec des noyaux Gaussiens d'écart-types croissants et à détecter les seuils délimitant les pics de chacun des histogrammes dans un nouvel espace appelé *scale-space*. Cet espace à deux dimensions admet les valeurs d'une composante couleur en abscisse et l'écart-type du noyau Gaussien  $g(x, \tau)$  en ordonnée. La plage d'échelles pour lesquelles ces seuils sont détectés dans le scale-space permet de distinguer les pics significatifs des artefacts de l'histogramme.

Dans cette section, nous détaillerons les différentes étapes nécessaires à cette méthode. La première d'entre-elles consiste à projeter des valeurs de seuils caractérisant les pics des histogrammes lissés dans le scale-space afin d'obtenir l'*empreinte* (fingerprint en anglais) de l'histogramme à analyser. Ensuite, nous expliquerons comment l'empreinte est analysée pour partitionner le scale-space en vue d'une sélection des pics significatifs de l'histogramme. La dernière partie traitera alors du critère utilisé pour la sélection des pics significatifs.

#### **3.5.2.1** Construction de l'empreinte

Dans la littérature, les seuils utilisés lors du SSF sont les points d'inflexion des courbes représentatives des histogrammes. Cependant, nous avons montré dans le paragraphe 3.3.3 que l'utilisation des minima locaux semble être plus judicieux pour segmenter une image couleur. Pour cette raison, nous utilisons les minima locaux comme seuils permettant de délimiter les pics des histogrammes pour le SSF. De même, les résultats présentés dans ce chapitre utilisent les minima locaux comme seuils délimitant les pics significatifs plutôt que les points d'inflexion.

L'analyse d'un histogramme lissé  $h^k(x,\tau)$  permet de déterminer des seuils  $S_{g,i}^k$  et  $S_{d,i}^k$ . Ces seuils permettent de construire des points de coordonnées  $(S_{g,i}^k,\tau)$  et  $(S_{d,i}^k,\tau)$  dans le scalespace. La figure 3.19 montre les coordonnées de tels points dans le scale-space à partir des histogrammes  $h^B(x,2)$ ,  $h^B(x,5)$  et  $h^B(x,10)$  lorsque les minima locaux sont utilisés comme seuils. Les points rouges sur l'axe des abscisses des figures 3.19(a), 3.19(b) et 3.19(c) renseignent sur la localisation des seuils détectés. Les valeurs obtenues par l'analyse des passages par zéro de la dérivé première de ces histogrammes sont portées dans le tableau 3.19(d) ainsi que les écarts-types du noyau Gaussien utilisé. Ce tableau renseigne par conséquent sur les coordonnées  $(S_{g,i}^k, \tau)$  et  $(S_{d,i}^k, \tau)$  correspondant à la projection des minima locaux dans le scale-space. On remarque que les seuils sont modifiés avec l'augmentation de l'échelle  $\tau$  et finissent par disparaître dans le cas ou des pics sont fusionnés. Les projections des seuils ainsi déterminés sont représentées graphiquement par des points rouges dans le scale-space sur la figure 3.20(a).





τ	$S_{g,1}^B$	$S_{d,1}^B$	$S^B_{g,2}$	$S^B_{d,2}$	$S^B_{g,3}$	$S^{B}_{d,3}$
2	0	72	72	105	105	255
5	0	73	/	/	73	255
10	0	78	/	/	78	255



**Figure 3.19 :** Détection des valeurs sur l'axe des abscisses des seuils des histogrammes  $h^B(x,2)$ ,  $h^B(x,5)$  et  $h^B(x,10)$  de l'image de synthèse. Les points en rouge montrent la localisation de ces seuils.

Les valeurs de seuil pour l'ensemble des échelles  $\tau$  analysées délimitant un même pic donnent naissance à une courbe dans le scale-space appelée *scale-space image*. L'ensemble des scale-space image d'un histogramme  $h^k(x)$  est appelé empreinte de l'histogramme. La figure 3.20(b) montre l'empreinte de l'hitogramme  $h^B(x)$  pour l'ensemble des échelles  $\tau$  tel que  $1 \le \tau \le 26$ .

Sur la figure 3.20(b), les scale-space images déterminés avec les minima locaux sont en rouge. Les scale-space image des maxima locaux des histogrammes ont été rajoutés en vert sur la figure afin de mieux appréhender l'allure du pic pour les différentes échelles. On peut ainsi remarquer que lors de l'apparition d'un pic, la coordonnée correspondant à un maximum local est proche la coordonnée correspondant à un minimum local. Plus le paramètre d'échelle diminue, plus la valeur d'abscisse à laquelle se situe l'amplitude maximale de ce pic s'éloigne de ce seuil. Enfin, on peut remarquer sur la figure 3.20(b) que pour les valeurs du paramètre d'échelle les plus importantes, il n'y a qu'un seul pic détecté et que celui-ci est défini sur l'intervalle [0,255].





(a) Construction de l'empreinte de l'histogramme (b) Empreinte de l'histogramme de la composante  $h^{B}(x)$  de la composante bleue de l'image de bleue  $h^{B}(x)$  de l'image de synthèse 3.3(a). synthèse 3.3(a) à partir des minima locaux détectés sur les histogrammes  $h^B(x,2)$ ,  $h^B(x,5)$  et  $h^{B}(x, 10).$ 

Figure 3.20 : Représentation des minima locaux des histogrammes lissés de la composante bleue de l'image de synthèse par des noyaux gaussiens d'écarts-type 2, 5 et 10 dans le scale-space et l'empreinte de l'histogramme  $h^B(x)$ .

Les figures 3.21(a), 3.21(b) et 3.21(c) montrent les représentations graphiques des empreintes calculées lorsque les valeurs de seuils sont les minima locaux pour les histogrammes des composantes rouge  $h^{R}(x)$ , verte  $h^{V}(x)$  et bleue  $h^{B}(x)$  de l'image de synthèse. Sur ces figures, les scale-space images correspondant aux minima locaux sont en rouge. Les scale-space images correspondant aux maxima locaux des pics ont été rajoutés en vert. La raison pour laquelle les plages d'échelles sont différentes pour les figures 3.21(a), 3.21(b) et 3.21(c) sera expliquée lors du paragraphe 3.5.3.4.



(a) Empreinte de l'histogramme (b) Empreinte de l'histogramme (c) Empreinte de l'histogramme  $h^{R}(x)$ .  $h^{V}(x)$ .  $h^{B}(x)$ .

Figure 3.21 : Empreintes des histogrammes 3.4(a) 3.4(b) et 3.4(c) de l'image 3.3(a).

D'un point de vue pratique, on ne calcule les histogrammes lissés qu'à partir d'une valeur minimale d'écart-type car la nature discrète des niveaux des composantes rend difficile l'obtention de l'empreinte dans les basses échelles. Ce seuil minimal n'est parfois pas précisé dans la littérature [Sch93, CL03] ou égal à 1 [LL90, Liu93, LY94].

#### 3.5.2.2 Construction des nœuds

L'analyse des empreintes ainsi déterminées dans le scale-space permet de représenter graphiquement toutes les partitions possibles de l'histogramme suivant les seuils détectés.

Pour la valeur d'écart-type maximale  $\tau_{max}$ , les seuils déterminés à partir de l'histogramme  $h^k(x, \tau_{max})$  partitionnent la droite d'équation  $\tau = \tau_{max}$  dans le scale-space en intervalles. Si l'on diminue  $\tau$  progressivement, des seuils délimitant des pics qui n'étaient pas présents à une échelle plus grande apparaissent, donnant ainsi naissance à des sous-intervalles. Chacun de ces sous-intervalles est inclus dans un rectangle dans le scale-space appelé *nœud*. Ces nœuds seront ensuite examinés pour déterminer ceux qui correspondent à des pics significatifs de l'histogramme étudié. Ces rectangles sont délimités :

- en haut par l'échelle à laquelle le sous-intervalle émerge,
- en bas par l'échelle à laquelle un autre sous-intervalle apparaît,
- à gauche et à droite par les valeurs d'abscisses des scale-space images qui sont les plus faibles.

La figure 3.22 montre les nœuds de l'histogramme  $h^B(x)$  construis à partir de ces règles. Sur cette figure, l'empreinte de la figure 3.20(b) a été superposée aux nœuds représentés par leurs rectangles.
Examinons les bornes délimitant le nœud 4 dans le scale-space représenté sur cette figure. Celui-ci est limité à gauche, à droite, en haut et en bas respectivement par les valeurs 72, 106, 3.8 et 1 sur cette figure 3.22. En effet, L'échelle à laquelle le sous-intervalle émerge dans le scalespace est  $\tau = 3.8$ , l'échelle à laquelle un autre sous-intervalle apparaît dans le scale-space est  $\tau = 1$ . La valeur d'abscisse la plus faible du scale-space image pour laquelle le sous-intervalle a émergé est x = 106 et la valeur d'abscisse du scale-space image de l'intervalle dont il est issu vaut x = 72.



**Figure 3.22 :** Nœuds de l'histogramme  $h^B(x)$ .

Les figures 3.23(a), 3.23(b) et 3.23(c) montrent respectivement les représentations des nœuds calculés pour les histogrammes des composantes rouge  $h^R(x)$ , verte  $h^V(x)$  et bleue  $h^B(x)$  de l'image de synthèse en utilisant les minima locaux pour détecter les seuils. Sur ces figures, les rectangles correspondant aux nœuds ont été superposés et numérotés en bleu aux empreintes des histogrammes.



(a) Noeuds de l'histogramme (b) Noeuds de l'histogramme (c) Noeuds de l'histogramme  $h^{R}(x)$ .  $h^{V}(x)$ .

**Figure 3.23 :** Noeuds des histogrammes des composantes couleur rouge 3.4(a), verte 3.4(b) et bleue 3.4(c) de l'image de synthèse 3.3(a).

Il se pose alors la question du critère à utiliser afin de sélectionner les nœuds correspondant à des pics significatifs de l'histogramme analysé. De tels nœuds sont aussi appelés *nœuds actifs*.

#### 3.5.2.3 Sélection des nœuds actifs

Le critère proposé par Witkin pour déterminer les nœuds actifs est basé sur la hauteur des rectangles associés aux nœuds suivant l'axe des ordonnées dans le scale-space.

Pour introduire ce critère, il nous faut définir la notion de *descendance* entre nœuds. Le SSF peut-être vu comme un arbre puisqu'un intervalle peut-être divisé en sous-intervalles lors de l'apparition d'un nouveau pic dans l'histogramme lissé à une échelle inférieure. Ces sous-intervalles donnent alors naissance à de nouveaux nœuds représentant la descendance. Le nœud situé au-dessus de la valeur de l'échelle à laquelle un nouveaux pic apparaît est appelé *nœud père* et les nœuds situés en-dessous de la valeur de l'échelle à laquelle un nouveaux pic apparaît sont appelés *nœud fils*. Par exemple, le nœud étiqueté 3 de la figure 3.23(c) est le nœud père des nœuds étiquetés 4 et 5. Ainsi, le critère utilisé par le SSF peut-être énoncé avec la formulation suivante :

Un nœud est actif si la hauteur du rectangle qui le définit, c'est-à-dire la distance entre sa limite inférieure et sa limite supérieure, est supérieure à la hauteur moyenne des rectangles de ces descendants.

La plage d'échelles pour laquelle les nœuds sont définis dans le scale-space est appelée *stabilité* ou *persistance*.

Ainsi, plus la plage d'échelles d'un nœud est importante, plus ce nœud sera considéré comme stable ou persistant. La figure 3.24 montre les nœuds actifs de l'histogramme  $h^B(x)$ 

représentés par des rectangles grisés.



**Figure 3.24 :** Nœuds actifs sélectionnés par le SSF de l'histogramme  $h^{B}(x)$ .

Un point important à souligner provient du choix de la valeur d'échelle maximale  $\tau_{max}$ permettant ainsi de définir la plage d'échelles à analyser. En effet, une échelle maximale  $\tau_{max}$ petite ne permet pas de sélectionner les pics qui sont les plus stables dans le scale-space puisque seules les basses échelles sont analysées. A l'inverse, une échelle maximale  $\tau_{max}$  trop grande ne permet pas de sélectionner automatiquement le nœud actif correspondant à celui défini pour la valeur d'échelle  $\tau_{max}$  puisque la stabilité de celui-ci est beaucoup plus grande que celle de ses descendants. Dans la littérature, cette valeur  $\tau_{max}$  est généralement fixée de manière empirique. Cependant, elle peut être ajustable par l'utilisateur permettant ainsi de régler la finesse de la segmentation souhaitée [LY94].

Rosin souligne que Witkin a expérimenté d'autres critères de sélection, mais ces derniers n'ont pas été spécifiés [Ros98]. Ainsi, le critère énoncé est, à notre connaissance, le seul utilisé pour l'analyse d'histogrammes en vue de la segmentation d'images couleur. Le paragraphe suivant met en évidence les limites atteintes par le SSF utilisant ce critère.

#### 3.5.2.4 Limites du critère de sélection des nœuds actifs

Afin de mettre en avant les limites atteintes par la méthode du scale-space filtering (SSF), nous proposons de vérifier si les pics correspondant aux nœuds actifs de l'histogramme  $h^B(x)$ discriminent effectivement les différentes régions de l'image de synthèse 3.3(a). Chaque nœud calculé à partir de l'histogramme  $h^B(x)$  est numéroté de 1 à 7 sur la figure 3.24. Les nœuds étiquetés 2 et 3 sont considérés comme actifs par le SSF. Les pixels de l'image de synthèse 3.3(a) dont les niveaux de bleu appartiennent au nœud actif 2 sont montrés sur l'image 3.25(a) et ceux appartenant au nœud actif 3 sont montrés sur l'image 3.25(b).



(a) Pixels de l'image de synthèse 3.3(a) dont le niveau de bleu appartiennent au pic correspondant au nœud actif 2.



(b) Pixels de l'image de synthèse 3.3(a) dont le niveau de bleu appartiennent au pic correspondant au nœud actif 3.

**Figure 3.25 :** Pixels de l'image de synthèse 3.3(a) dont les niveaux de bleu appartiennent aux pics correspondant aux nœuds actifs sélectionnés par le critère clasique du SSF.

On peut remarquer que le premier pic correspondant au nœud 2 permet d'extraire le fond marron R1, le carré orange R2 et le carré jaune R4. Le deuxième pic permet quant à lui d'isoler l'anneau vert clair R5 le disque vert foncé R6 et la tache mauve R3.

Puis, dans un deuxième temps, nous avons sélectionné de manière interactive d'autres nœuds actifs. Cette nouvelle sélection a pour objectif de montrer qu'elle permet de mieux discriminer les régions de l'image que celle issue du critère utilisé par le SSF. Les nœuds actifs ainsi sélectionnés sont grisés sur la figure 3.26 et sont numérotés 2, 4 et 5.



**Figure 3.26 :** Nœuds actifs de l'histogramme  $h^{B}(x)$  sélectionnés interactivement.

Les pixels de l'image de synthèse 3.3(a) dont les niveaux de bleu appartiennent au nœud actif 2 sont montrés sur l'image 3.27(a), ceux appartenant au pic correspondant au nœud actif 4 sont montrés sur l'image 3.27(b) et ceux appartenant au pic correspondant au nœud actif 5 sont montrés sur l'image 3.27(c).



le niveau de bleu appartiennent le niveau de bleu appartiennent le niveau de bleu appartiennent au pic correspondant au nœud actif 2.

au pic correspondant au nœud au pic correspondant au nœud actif 4.

actif 5.

Figure 3.27 : Pixels de l'image de synthèse 3.3(a) dont les niveaux de bleu appartiennent aux pics correspondant aux nœuds actifs sélectionnés de manière interactive.

On remarque que cette sélection interactive des nœuds actifs permet de discriminer la région

de l'anneau vert clair  $\Re 5$  du disque vert foncé  $\Re 6$ , contrairement aux nœuds actifs déterminés par le SSF. Cet exemple montre ainsi que le critère utilisé par le SSF pour déterminer les nœuds actifs ne fournit pas toujours les meilleurs résultats possibles en termes de classification de pixels. Nous proposons par conséquent d'utiliser un critère de sélection des nœuds actifs qui prend en compte l'arrangement spatial des pixels dans l'image plutôt que la stabilité des nœuds dans le scale-space. Cette méthode basée sur ce nouveau critère sera appelée *Spatial Scale-space Filtering*.

#### 3.5.3 Spatial Scale-space Filtering

Certains pics correspondant à de petites régions de l'image donnent naissance à des pics de faibles amplitudes qui sont assimilés à des artefacts par le critère utilisé par le SSF puisqu'ils ne sont pas stables dans le scale-space. Nous proposons dans cette section un nouveau critère de sélection des nœuds actifs basé sur l'analyse spatiale des pixels [BVMP06]. Ce critère utilise le degré de compacité proposé par Macaire défini comme le produit du degré d'homogénéité et du degré de connexité [Mac04, MVP06]. Ainsi, avant d'expliquer le critère proposé nous allons définir le degré de connexité, le degré d'homogénéité et le degré de compacité.

#### 3.5.3.1 Le degré de connexité

On note  $S_i^k$  le sous-ensemble constitué de tous les pixels P d'une image I dont les niveaux  $C^k(P)$  appartiennent à l'intervalle délimitant le  $i^{\text{ème}}$  pic de la composante couleur k défini par  $[S_{g,i}^k, S_{d,i}^k]$ .

Pour chaque pixel P de  $S_i^k$ , on définit  $N_{S_i^k}(P)$ , le sous-ensemble des pixels 8 voisins Q de P qui appartiennent également au sous-ensemble  $S_i^k$ . La connexité entre P et  $S_i^k$  noté  $\gamma_{S_i^k}(P)$  dépend du nombre de pixels Q qui appartiennent au sous-ensemble  $N_{S_i^k}(P)$  vérifiant la relation 3.4 :

$$\gamma_{S_i^k}(P) = \frac{card\left\{Q \in N_{S_i^k}(P)\right\}}{8}.$$
(3.4)

Le degré de connexité  $CD(S_i^k)$  du sous-ensemble  $S_i^k$  se définit alors par la relation 3.5 :

$$CD(S_i^k) = \frac{\sum\limits_{P \in S_i^k} \gamma_{S_i^k}(P)}{card\left\{P \in S_i^k\right\}}.$$
(3.5)

Le degré de connexité vaut 0 pour un sous-ensemble de pixels vide. Un degré de connexité faible (proche de 0) correspond à un sous-ensemble de pixels dispersés dans l'image tandis qu'un degré de connexité fort (proche de 1) signifie que les pixels du sous-ensemble sont voisins les uns des autres dans l'image et correspondent à une ou plusieurs régions dans l'image de la composante considérée.

Pour distinguer le cas où  $S_i^k$  correspond à une seule région de l'image du cas où il correspond à plusieurs régions, nous proposons d'introduire le degré d'homogénéité [MVP04].

#### 3.5.3.2 Le degré d'homogénéité

Le degré d'homogénéité noté  $DH(S_i^k)$  se base sur la mesure de dispersion  $\sigma(S_i^k)$  des niveaux des pixels appartenant au sous-ensemble  $S_i^k$  à partir de la formule 3.6 :

$$\sigma(S_i^k) = \frac{1}{card\left\{P \in S_i^k\right\}} \times \sqrt{\sum_{P \in S_i^k} \left(C^k(P) - M^k(S_i^k)\right)^2}.$$
(3.6)

Dans cette formule  $M^k(S_i^k)$  est la valeur moyenne des niveaux  $C^k(P)$  des pixels P appartenant au sous-ensemble  $S_i^k$ .

Afin de décider si un sous-ensemble correspond effectivement à une région de l'image, la mesure globale de la dispersion des niveaux des pixels de ce sous-ensemble  $S_i^k$  est comparée avec une mesure locale de la dispersion des pixels voisins qui appartiennent au sous-ensemble.

La mesure locale de dispersion  $\sigma_{local}(S_i^k)$  des niveaux des pixels P est définie comme la moyenne des mesures de dispersion  $\sigma(N_{S_i^k}(P))$  évaluées aux voisinages  $N_{S_i^k}(P)$  de tous les pixels P. Cette mesure locale de dispersion pour le sous-ensemble  $S_i^k$  s'exprime par la formule 3.7 :

$$\sigma_{local}(S_i^k) = \frac{1}{card\left\{P \in S_i^k\right\}} \times \sum_{P \in S_i^k} \sigma(N_{S_i^k}(P))$$
(3.7)

Si la mesure locale de dispersion  $\sigma_{local}(S_i^k)$  des niveaux des pixels *P* est proche de la mesure globale  $\sigma(S_i^k)$ , les niveaux des pixels de ce sous-ensemble ont tendance à donner naissance à un pic unique dans l'histogramme. Par contre, si  $\sigma_{local}(S_i^k)$  est inférieure à  $\sigma(S_i^k)$ , les niveaux de couleurs des pixels tendent à être répartis en plusieurs pics séparés dans l'histogramme de la composante *k*.

Afin de comparer  $\sigma_{local}(S_i^k)$  et  $\sigma(S_i^k)$ , le degré d'homogénéité  $DH(S_i^k)$  du sous-ensemble  $S_i^k$  est défini par la relation 3.8 :

$$\begin{cases} DH(S_i^k) = 1 & \text{si} \quad \sigma(S_i^k) = 0, \\ DH(S_i^k) = \frac{\sigma_{local}(S_i^k)}{\sigma(S_i^k)} & \text{si} \quad \sigma(S_i^k) > 0. \end{cases}$$
(3.8)

Le degré d'homogénéité  $DH(S_i^k)$  varie entre 0 et 1. Afin d'éviter que le degré d'homogénéité  $DH(S_i^k)$  ne soit supérieur à 1, la valeur  $DH(S_i^k)$  est fixée à 1 si  $\sigma_{local}(S_i^k)$  est supérieur à  $\sigma(S_i^k)$ . Quand les niveaux  $C^k(P)$  des pixels considérés forment un pic compact dans l'histogramme de la composante couleur k, alors le degré d'homogénéité est proche de 1. A l'inverse, si les niveaux des pixels constituent plusieurs pics bien séparés, le degré d'homogénéité est proche de 0.

Le degré de connexité et le degré d'homogénéité d'un sous-ensemble fournissent des informations complémentaires sur les pixels appartenant au sous-ensemble étudié. Le degré de connexité renseigne si les pixels appartenant au sous-ensemble sont voisins les uns des autres tandis que le degré d'homogénéité renseigne si les niveaux des pixels appartenant au sousensemble forment un pic compact ou non. Par conséquent, l'analyse conjointe du degré de connexité et du degré d'homogénéité d'un sous-ensemble mérite d'être abordée pour estimer le degré de confiance à accorder à la proposition : "le sous-ensemble considéré correspond effectivement à une région de l'image".

#### 3.5.3.3 Le degré de compacité

Afin de tenir compte simultanément des propriétés de connexité et d'homogénéité d'un sous-ensemble, le degré de compacité noté  $DCH(S_i^k)$  est défini par la relation 3.9 :

$$DCH(S_i^k) = CD(S_i^k) \times DH(S_i^k)$$
(3.9)

Le degré de compacité est proche de 1 si les pixels du sous-ensemble sont fortement connexes dans l'image  $(DC(S_i^k) \approx 1)$  et si les niveaux des pixels forment un pic compact dans l'histogramme  $(DH(S_i^k) \approx 1)$ , tandis qu'un degré de compacité proche de 0 signifie que les pixels du sous-ensemble sont dispersés dans l'image  $(DC(S_i^k) \approx 0)$  et/ou que les niveaux des pixels ne forment pas un unique pic compact dans l'histogramme  $(DH(S_i^k) \approx 0)$ .

#### 3.5.3.4 Définition du critère de sélection des nœuds actifs

Chaque nœud du scale-space peut-être associé à un sous-ensemble de pixels  $S_i^k$ . Nous proposons par conséquent de mesurer les degrés de compacité des sous-ensembles associés aux nœuds. A titre d'exemple, la figure 3.28 renseigne sur les degrés de compacité des sous-ensembles de pixels associés aux nœuds de l'histogramme  $h^B(x)$  de l'image de synthèse. Les numéros des nœuds sont écrit entre parenthèses dans la figure. Cette figure montre que lorsque le paramètre d'échelle  $\tau$  diminue, le nombre N de nœuds augmente parce qu'un nœud père donnera naissance à au moins deux nœuds fils pour une valeur spécifique d'échelle  $\tau$ . Les nouveaux degré de compacité des nœuds fils sont alors comparés avec celui du nœud père. Si au moins un des degrés de compacité des nœuds fils est inférieur au degré de compacité du nœud père, alors les nœuds fils ne peuvent pas être sélectionnés pour être des nœuds actifs. Le critère proposé est défini par la règle suivante :

Un nœud est actif si le degré de compacité d'au moins un de ses nœuds fils est inférieur à celui du nœud père considéré.



Figure 3.28 : Degré de compacité des noeuds provenant de l'histogramme  $l^{\mathcal{B}}(x)$  de l'image de synthèse 3.3(a).

Par exemple, le degré de compacité du nœud 3 vaut 0.526. Ce nœud se divise en deux nœuds fils 4 et 5 dont les degrés de compacité valent respectivement 0.719 et 0.793. Ces deux valeurs du degré de compacité sont supérieures à celle du nœud père. Cela signifie que les pixels associés aux deux nœuds fils représentent de manière plus probable des régions dans l'image que les pixels associés au nœud 3. Dans cet exemple, les nœuds actifs sont les nœuds 2, 4 et 5.

Tandis que le critère utilisé par la méthode du SSF est basé sur la stabilité des nœuds dans le scale-space, le nouveau critère proposé est basé sur les degrés de compacité. Ce nouveau critère proposé permet de s'affranchir par conséquent de la plage d'échelles à analyser pour déterminer les nœuds actifs contrairement au critère traditionnellement utilisé. En effet, nous avons montré dans le paragraphe 3.5.2.3 que la sélection des nœuds actifs par le SSF dépend de l'échelle maximale  $\tau_{max}$  pour laquelle l'empreinte est analysée. Avec le nouveau critère proposé, le spatial scale-space filtering balaie l'ensemble des échelles jusqu'à ce que l'histogramme lissé ne possède qu'un seul pic pour avoir une étude complète de son empreinte.

#### 3.5.4 Résultats de l'image de synthèse

Afin de mesurer l'apport du critère proposé, nous proposons de comparer les résultats de segmentation fournis par le SSF et par le spatial scale-space filtering. La figure 3.29 montrent les résultats ainsi obtenus avec l'image de synthèse 3.3(a).

Avec le critère de sélection des nœuds actifs utilisé par le SSF basé sur la stabilité des nœuds, les histogrammes lissés  $h^R(x, \tau)$ ,  $h^V(x, \tau)$  et  $h^B(x, \tau)$  possèdent respectivement  $N_R = 2$ ,  $N_V = 2$  et  $N_B = 2$  pics et permettent de partitionner l'espace couleur (R, V, B) uniquement en  $N_R \times N_V \times N_B = 8$  parallélépipèdes rectangles définissant ainsi huit classes de pixels possibles. Parmi ces 8 classes possibles, seulement  $N_C = 4$  d'entre elles admettent des populations de pixels. L'image 3.29(a) montre les pixels assignés aux 4 classes.

Le nombre de pics significatifs des histogrammes lissés  $h^R(x, \tau)$ ,  $h^V(x, \tau)$  et  $h^B(x, \tau)$  détectés par le spatial SSF sont  $N_R = 2$ ,  $N_V = 4$  et  $N_B = 3$ . Ceci permet de partitionner l'espace couleur (R, V, B) en  $N_R \times N_V \times N_B = 24$  parallélépipèdes, définissant 24 classes possibles. Parmi ces 24 classes possibles,  $N_C = 6$  d'entre elles admettent des populations de pixels. L'image 3.29(b) montre les pixels assignés à chacune de ces 6 classes.



(a) Image segmentée en utilisant le SSF.



(b) Image segmentée en utilisant le spatial scalespace filtering.

**Figure 3.29 :** Résultats de la segmentation de l'image de synthèse 3.3(a) avec deux critères de sélection des nœuds actifs différents.

Les résultats montrent que le SSF ne permet pas de séparer la région  $\Re 5$  de la  $\Re 6$  (l'anneau vert clair et le disque vert foncé). Par ailleurs, les pixels des régions  $\Re 2$  et  $\Re 4$  (le carré orange et le carré jaune) sont assignés à la même classe de pixels (c.f. figure 3.29(a)).

L'image 3.29(b) montre les résultats obtenus quand les pics sont détectés par le spatial scalespace filtering. Les 6 classes construites correspondent effectivement aux 6 régions de l'image, et seuls quelques pixels entre les régions  $\Re 5$  et  $\Re 6$  ont été mal classés à cause du chevauchement important des distributions de points couleur correspondant à ces deux régions dans l'espace (R, V, B).

# 3.6 Résultats expérimentaux

Dans ce paragraphe, nous présentons des résultats obtenus par notre méthode appliquée à la segmentation d'images de référence. Ainsi les seuils le long des trois composantes R, V et B sont détectés par le spatial scale-space filtering appliqué aux histogrammes 1D de ces composantes couleur.

L'image "house" de la figure 3.30(a) présente plusieurs difficultés en vue d'une segmentation à cause des effets d'ombre sous le toit de la maison. Par ailleurs, le mur et le toit présentent des couleur non-homogènes. Afin d'illustrer le résultat de la segmentation obtenue, l'image résultat de la figure 3.30(b) est affichée en fausses couleurs. L'image segmentée est constituée de



(a) Image originale "house".

(b) Résultat l'image "house" segmentée.

Figure 3.30 : Image originale et image segmentée de l'image "house".

six classes correspondant au ciel, aux murs, aux fenêtres, aux ombres, au toit et aux gouttières.

L'image "jelly beans" de la figure 3.31(a) est constituée de bonbons de quatre couleurs différentes (rouge, vert, jaune et marron) présentant de nombreux ombres et reflets. On peut aussi remarquer que les intensités des pixels représentant le fond de l'image ne sont pas homogènes. Le résultat de la segmentation fournit cinq classes de pixels correspondant aux quatre types de bonbons et au fond (cf. image de la figure 3.31(b)). Malgré la non-uniformité de l'intensité des pixels caractérisant le fond, ces pixels sont assignés à une seule classe. Par ailleurs, les pixels représentant les bonbons sont correctement classés.

Visuellement, l'image "hand" de la figure 3.32(a) est constituée de trois classes de pixels correspondant au fond, la main et l'anneau bleu. Il n'est pas facile de segmenter cette image correctement car les points couleur dans l'espace (R, V, B) des pixels de ces trois classes forment des nuages de points qui se chevauchent. Le résultat de la segmentation de la figure 3.32(b) est cependant satisfaisant car trois classes ont été construites. Cependant, nous constatons qu'une partie des pixels de l'anneau est assignée à la classe du fond. Nous constatons également que les



(a) Image originale "jelly beans".

(b) Résultat l'image "jelly beans" segmentée.





- (a) Image originale "hand".
- (b) Résultat l'image "hand" segmentée.
- Figure 3.32 : Image originale et image segmentée de l'image "hand".

pixels du fond sont assignés à une seule classe de pixels alors que celui-ci est texturé. De plus, l'information des ongles est conservée car les pixels représentant les ongles sont regroupés en une classe.

# 3.7 Conclusion

Dans ce chapitre, nous avons présenté une méthode de segmentation par multi-seuillage des composantes couleur où les couleurs des pixels sont codées uniquement dans l'espace (R, V, B). Cette méthode recherche les seuils délimitant les pics des histogrammes 1*D* des composantes couleur. Lorsque sont retrouvés les intervalles représentant les pics de chaque histogramme 1*D*, l'espace couleur est partitionné en parallélépipèdes rectangles grâce au produit cartésien de ces

intervalles. Une analyse de la population des pixels dont les points couleur se trouvent dans ces parallélépipèdes permet d'identifier les domaines de définition des classes. Une fois les pixels assignés aux classes ainsi construites, une analyse en composantes connexes est effectuée pour produire une image segmentée en régions.

Le point clé de cette approche est la détection des seuils le long de chaque composante couleur. Pour les détecter, il est souvent nécessaire de lisser l'histogramme 1*D* au préalable. La méthodes de lissage proposée par Lin et al. [LWY96], basée sur des hypothèses heuristiques, se révèle être rapide puisqu'elle lisse les histogrammes avec un nombre faible d'itérations (3 ou 4 généralement). Cependant, des résultats expérimentaux ont montré que cette méthode a tendance à éliminer des pics significatifs. Par conséquent, certains pics significatifs des histogrammes correspondant à des régions de couleurs différentes sont considérés comme formant un seul pic. L'utilisation de cette méthode a donc tendance à sous-segmenter l'image.

Nous avons décrit en détail la méthode dite "scale-space filtering" qui analyse la réponse de l'histogramme à une batterie de filtres Gaussiens pour détecter les seuils délimitant les pics. Les réponses classées par valeurs d'écart-type croissantes constituent le scale space. Une analyse des scale space images reconstruit des nœuds qui correspondent aux pics candidats. Le critère de sélection des nœuds à retenir utilisé par le scale-space filtering ne tient compte que de la taille du nœud dans le scale space. Nous avons alors montré les limites atteintes par une telle approche et avons proposé un nouveau critère de sélection des nœuds.

Ce critère qui est exploité par le spatial scale space filtering tient compte de la répartition spatiale des pixels dans l'image dont les niveaux de la composante couleur sont associés à chaque nœud. Des résultats expérimentaux effectués sur une image de synthèse ont montré l'amélioration en termes de taux de bonne classification apportée par le spatial scale space filtering. Cette méthode a également été testée sur des images de référence dont la couleur est codée dans l'espace (R,V,B). Les résultats obtenus sont intéressants. Toutefois il est pertinent d'examiner si le codage des couleurs dans d'autres espaces couleur ne permet pas d'améliorer les résultats de segmentation.

Le choix de l'espace couleur en vue de segmenter une image couleur est un problème qui a été souvent abordé dans la littérature. Les nombreuses études menées sur ce thème n'ont pas permis de dégager un espace couleur qui permet d'obtenir une segmentation pertinente de tous les types d'images. Le prochain chapitre traitera de notre contribution à cette étude en présentant une méthode de segmentation d'images couleur dont la principale originalité consiste à déterminer l'espace couleur adapté à la construction de chaque classe de pixels.

# **Chapitre 4**

# Sélection automatique d'espaces couleur pour la segmentation d'images

Dans le chapitre précédent, nous avons présenté une méthode de segmentation d'images couleur par multi-seuillage sur chacune des composantes couleur. Nous avons vu que la sélection des seuils est l'un des points clés de cette approche. Ohlander tente de réduire l'influence de la sélection des seuils en proposant une analyse itérative des histogrammes mono-dimensionnels afin de construire les classes par multi-seuillages successifs [OPR78]. Il a montré que cette approche permet de mettre en évidence des pics significatifs qui ne sont pas distinguables en prenant en considération l'ensemble des pixels de l'image. Nous proposons de nous inspirer de cette approche afin de construire les classes de manière itérative.

Par ailleurs, la construction des classes par la méthode décrite dans le chapitre précédent ne s'opère que dans l'espace couleur (R, V, B) alors que Vandenbroucke a montré que le choix de l'espace de représentation de la couleur conditionne la qualité de la segmentation [VMP03]. Pour cette raison, nous proposons de sélectionner de manière non supervisée l'espace couleur adapté qui met en évidence chacune des classes à construire [BVMP04a, BVMP04b, BVMP05].

Le premier paragraphe de ce chapitre est consacré à la présentation du principe de l'algorithme proposé qui se divise en six étapes successives. Cet algorithme est itératif et s'arrête quand un critère d'arrêt n'est pas vérifié. Dans les paragraphes suivants, nous détaillons chacune de ces étapes. Le dernier paragraphe recense quelques résultats obtenus par notre méthode. Pour conclure, nous menons une synthèse des résultats obtenus.

# 4.1 Méthode de segmentation proposée

La méthode de segmentation que nous proposons consiste en un multi-seuillage par analyse itérative d'histogrammes mono-dimensionnels. L'originalité principale de la méthode proposée consiste en la sélection automatique de l'espace couleur pour mettre en évidence chacune des classes en présence. La figure 4.1 montre les différentes étapes de l'algorithme itératif qui sont détaillées dans les paragraphes suivants. L'algorithme se divise en 6 étapes successives qui sont réitérées tant qu'un critère d'arrêt n'est pas vérifié.



Figure 4.1 : Algorithme de segmentation d'images couleur proposé par construction itérative des classes de pixels.

La première étape consiste à coder les couleurs des pixels dans différents espaces couleur et de calculer les histogrammes 1*D* des composantes couleur des différents espaces considérés. Ces histogrammes sont filtrés lors de la deuxième étape par une méthode de lissage afin de mettre en évidence les pics significatifs. Les seuils délimitant ces pics sont détectés lors de la troisième étape. A la quatrième étape, le pouvoir discriminant de chacun des espaces couleur est calculé à partir des seuils détectés par l'analyse des histogrammes. L'espace couleur le plus discriminant est sélectionné parmi les espaces disponibles afin de construire une classe de pixels lors de la cinquième étape. Les pixels assignés à cette classe sont alors extraits de l'image à analyser durant la sixième étape. L'algorithme est réitéré avec les pixels qui n'ont été assignés à aucune classe tant qu'un critère d'arrêt basé sur le nombre de pixels assignés n'est pas vérifié.

# **4.2 Détermination des histogrammes** 1D

A chaque itération, la couleur des pixels de l'image à segmenter est codée dans chacun des  $N_S$  espaces de la représentation de la couleur ( $N_S = 11$ ) suivants :

- 1. (R, V, B),
- 2. (r, v, b),
- 3. (X, Y, Z),
- 4. (x, y, z),
- 5. (Y, I, Q),
- 6. (Y, U, V),
- 7. (*I*1,*I*2,*I*3),
- 8.  $(A, C_1, C_2)$ ,
- 9.  $(Y, Ch_1, Ch_2)$ ,
- 10.  $(L^*, a^*, b^*)$ ,
- 11.  $(L^*, u^*, v^*)$ .

Ces espaces couleur ont été présentés dans le chapitre 2 et appartiennent à trois familles différentes d'espaces couleur suivant le regroupement établi par Vandenbroucke [Van00]. Les espaces utilisés par la méthode proposée sont indiqués en rouge sur la figure 4.2. On peut remarquer sur cette figure que les espaces couleur perceptuels ne sont pas pris en compte par notre méthode. Ces espaces couleur n'ont pas été retenus car ils possèdent une composante de teinte dont les niveaux son codés par un angle. En conséquence, la teinte ne peut pas être considérée comme les autres composantes couleur. De plus, les histogrammes des composantes des espaces perceptuels sont très bruités et par conséquent difficiles à exploiter pour en extraire les pics significatifs [Ken76].

La première étape de l'algorithme de la figure 4.1 consiste à calculer les histogrammes des composantes couleur de chacun de ces  $N_S$  espaces couleur. Dans le  $j^{\text{ème}}$  espace couleur avec  $1 \le j \le N_S$ , nous déterminons l'histogramme 1D  $h^{j,k}(x)$  de la  $k^{\text{ème}}$  composante couleur avec  $k \in \{1,2,3\}$ , où x est le niveau de couleur de la composante couleur étudiée. La figure 4.3 montre les histogrammes des trois composantes de l'espace (R,V,B) de l'image originale "hand". Cette image nous servira d'illustration tout au long du chapitre car elle fournit des résultats intéressants pour expliquer et justifier chacune des étapes proposées. La figure 4.3 montre les histogrammes des composantes rouge (c.f. figure 4.3(a)), verte (c.f. figure 4.3(b)) et bleue (c.f. figure 4.3(c)) de l'image "hand" (c.f. figure 4.3(d)).



Figure 4.2 : Les espaces couleur utilisés dans l'approche proposé sont notés en rouge.



**Figure 4.3 :** Histogrammes  $h^{1,1}(x)$ ,  $h^{1,2}(x)$  et  $h^{1,3}(x)$  des composantes rouge, verte et bleue de l'image "hand".

Nous avons vu au chapitre précédent qu'il est parfois délicat de déterminer les pics significatifs de tels histogrammes (cf. figures 4.3(a), 4.3(b) et 4.3(c)). Aussi, afin de révéler ces pics, la deuxième étape de la méthode proposée consiste en un lissage de ces histogrammes.

# 4.3 Lissage des histogrammes

Les résultats de la segmentation par multi-seuillage dépendent fortement de la qualité de lissage des histogrammes des composantes et de la détection des pics significatifs. Par ailleurs, nous souhaitons une procédure de lissage non-supervisée et non-paramétrique afin que la méthode fournisse des résultats satisfaisants pour tout type d'images.

Nous n'avons pas eu le temps d'intégrer le "spatial scale-space filtering" présenté au chapitre précédent à l'algorithme de segmentation que nous proposons dans ce chapitre. Par conséquent, chacun des histogrammes est lissé de manière adaptative en utilisant la méthode heuristique proposée par Lin et décrite au paragraphe 3.5.1 [LWY96]. L'histogramme lissé noté  $h^{j,k}(x, \tau)$  pour la  $k^{\text{ème}}$  composante du  $j^{\text{ème}}$  espace couleur est obtenu en effectuant le produit de convo-

lution entre l'histogramme 1*D* à traiter et un noyau Gaussien  $g(x, \tau)$  La méthode de Lin fixe automatiquement la valeur de l'écart-type  $\tau$  du noyau Gaussien de telle sorte que seuls les pics significatifs soient présents dans l'histogramme  $h^{j,k}(x, \tau)$ .

La figure 4.4 montre les histogrammes des composantes couleur R (c.f figure 4.4(a)), V (c.f figure 4.4(b)) et B (c.f figure 4.4(c)) de l'image "hand" ainsi que les histogrammes lissés de ces mêmes composantes couleur sur respectivement les figures 4.4(d), 4.4(e) et 4.4(f).



Figure 4.4 : Histogrammes lissés des composantes rouge, verte et bleue de l'image "hand".

Dans l'étape n°2 de l'algorithme de la figure 4.1, les histogrammes de chacune des composantes des espaces considérés sont ainsi lissés. Ces histogrammes sont ensuite analysés afin de déterminer les seuils délimitant les différents pics. Ces seuils serviront, par la suite, à la construction des classes de pixels en vue de segmenter l'image.

## 4.4 Détection des seuils

Nous avons vu dans le paragraphe 3.3, qu'il semble judicieux de détecter les minima locaux des histogrammes des composantes afin de déterminer les seuils utilisés pour le multi-seuillage d'une image couleur. Ces seuils sont donc déterminés par l'analyse des passages par zéro de l'estimation des dérivées premières des histogrammes. Le nombre de pics délimités par les seuils détectés par l'analyse de l'histogramme  $h^{j,k}(x,\tau)$  de la  $k^{\text{ème}}$  composante couleur du  $j^{\text{ème}}$  espace couleur est noté  $N_p^{j,k}$ .

Afin de ne prendre en compte que les pics les plus significatifs de l'histogramme  $h^{j,k}(x,\tau)$ , seuls ceux dont l'amplitude maximale est supérieure à 10% de l'amplitude maximale du pic

ayant la plus grande amplitude sont pris en compte lors de cette étape. Le  $i^{\text{ème}}$  pic significatif avec  $(1 \le i \le N_p^{j,k})$  de l'histogramme 1D  $h^{j,k}(x,\tau)$  est alors caractérisé par quatre paramètres :

- Le seuil gauche  $T_{g,i}^{j,k}$  le seuil droit  $T_{d,i}^{j,k}$ ,

- l'amplitude maximale 
$$H_i^{j,k} = \prod_{\substack{x \le T_{d,i}^{j,k} \\ x = T_{g,i}^{j,k}}} h^{j,k}(x,\tau)$$

- la valeur d'abscisse  $A_i^{j,k}$  correspondant à l'amplitude maximale  $H_i^{j,k}$ .

Les seuils gauche  $T_{g,i}^{j,k}$  et droit  $T_{d,i}^{j,k}$  permettent aussi de définir des sous-ensembles  $S_i^{j,k}$  de pixels dont les niveaux de la composante *k* appartiennent aux intervalles  $\left[T_{g,i}^{j,k}, T_{d,i}^{j,k}\right]$ .

De plus, sont déterminés pour chaque histogramme  $h^{j,k}(x,\tau)$ :

- le numéro  $P^{j,k}$  du pic ayant la plus grande amplitude :

$$P^{j,k} = \operatorname*{argmax}_{i \in \left[1, N_P^{j,k}\right]} (H_i^{j,k}(x, \tau))$$

$$(4.1)$$

– la plus petite distance  $D_P^{j,k}$  séparant deux pics significatifs voisins, à savoir la distance séparant les valeurs d'abscisse correspondant aux amplitudes maximales de deux pics voisins. Cette distance s'exprime par :

$$D_P^{j,k} = \min(\left|A_m^{j,k} - A_n^{j,k}\right|) \quad \text{avec} \quad m \neq n \text{ et } m \le N_P^{j,k}, n \le N_P^{j,k}.$$
(4.2)

La figure 4.5 montre les valeurs des différents paramètres calculées pour un histogramme construit à partir de trois Gaussiennes non-équiprobables centrées en 64, 128 et 192 et d'écarttype égal à 13. Notre algorithme ne détecte que deux pics significatifs parmi les trois. En effet, l'amplitude maximale du pic du milieu est inférieure à 10% de l'amplitude maximale du premier pic significatif. Ce pic n'est par conséquent pas considéré comme un pic significatif. Les valeurs  $T_{g,i}^{j,k}$ ,  $T_{d,i}^{j,k}$ ,  $A_i^{j,k}$  et  $H_i^{j,k}$  pour i = 1 et i = 2 sont reportées sur la figure ainsi que la distance  $D_P^{j,k}$ .

Après avoir détecté les seuils délimitant les pics significatifs présents dans les histogrammes des composantes des différents espaces couleur, l'étape n°4 de l'algorithme de la figure 4.1 recherche les pics qui correspondent le mieux aux différentes régions dans l'image. Cette recherche va être utilisée pour déterminer parmi les  $N_S$  espaces couleurs définis dans le paragraphe 4.2, celui qui permet la meilleure discrimination entre les classes en présence.

#### Sélection de l'espace couleur le plus discriminant 4.5

Un algorithme de segmentation d'images couleur fournit des résultats différents suivant l'espace de représentation de la couleur utilisé. Par ailleurs, de nombreux auteurs ont constaté qu'aucun espace couleur n'est mieux adapté qu'un autre pour segmenter tout type d'images [GH92, Van00, CJSW01]. Ainsi, l'utilisation d'un espace couleur particulier peut fournir de bons résultats pour des images acquises dans le cadre d'une application spécifique mais donner de mauvais



Figure 4.5 : Paramètres retournés par la méthode de détection des pics significatifs.

résultats sur des images d'une autre application. Vandebroucke a proposé de sélectionner automatiquement l'espace couleur adapté à la segmentation d'une image couleur en vue d'obtenir les meilleurs résultats possibles [Van00]. Cette sélection nécessite une phase d'apprentissage supervisé. Nous proposons de prolonger cette démarche en déterminant de manière non supervisée l'espace couleur le plus discriminant, et ce, à chaque itération de l'algorithme.

Pour cela, nous définissons le *pouvoir discriminant* d'un espace couleur comme étant sa capacité à distinguer les différentes classes de pixels représentant les régions effectives de l'image. Le pouvoir discriminant d'un espace sera mesuré en considérant de manière indépendante chacune de ses trois composantes. Nous proposons d'évaluer le pouvoir discriminant  $R^{j,k}$  de la  $k^{\text{ème}}$ composante du  $j^{\text{ème}}$  espace couleur à partir de l'analyse des  $N_P^{j,k}$  pics significatifs de l'histogramme  $h^{j,k}(x,\tau)$ . Pour ce faire, nous proposons un critère analysant l'arrangement spatial du sous-ensemble de pixels  $S_i^{j,k}$  associé à chaque pic significatif détecté.

#### 4.5.1 Pouvoir discriminant basé sur la répartition spatiale des pixels

Nous supposons que le pouvoir discriminant d'une composante couleur dépend conjointement du nombre de pics significatifs détectés par l'analyse de son histogramme et l'arrangement spatial des pixels dont les niveaux appartiennent à ces pics.

En effet, si un nombre important de pics significatifs permet de n'omettre aucune classe de pixels, mesurer l'arrangement spatial des pixels appartenant à ces pics permet d'estimer à quel point ces pics correspondent à des régions effectives de l'image. Nous proposons d'utiliser les degrés de connexité présentés dans le paragraphe 3.5.3.1 pour estimer l'arrangement spatial des pixels. Le degré de connexité d'un sous-ensemble  $S_i^{j,k}$  de pixels est proche de 1 quand ils forment des régions dans l'image tandis que le degré de connexité d'un sous-ensemble  $S_i^{j,k}$  de pixels est proche de 0 quand les pixels sont dispersés dans l'image.

Par conséquent, nous mesurons le pouvoir discriminant  $R^{j,k}$  d'une composante k d'un espace couleur j à partir de la somme des degrés de connexité  $CD(S_i^{j,k})$  des  $N_P^{j,k}$  sous-ensembles  $S_i^{j,k}$ associés aux pics significatifs :

$$R^{j,k} = \sum_{i=1}^{N_P^{j,k}} CD(S_i^{j,k})$$
(4.3)

Ce critère prend en compte simultanément le nombre de pics significatifs d'un histogramme et l'arrangement spatial des pixels de l'image dont les niveaux sont compris entre les seuils délimitant ces pics. En effet, plus l'histogramme  $h^{j,k}(x,\tau)$  de la  $j^{\text{ème}}$  composante couleur du  $k^{\text{ème}}$  espace couleur possède de pics significatifs, plus le nombre de termes pris en compte pour calculer la somme de l'équation 4.3 sera important. De même, plus les pixels formant les pics significatifs sont connexes, plus la somme des degrés de connexité des sous-ensembles correspondants sera élevée.

Afin d'illustrer le calcul du pouvoir discriminant des composantes couleur de l'espace (R, V, B) (j = 1), le tableau 4.1 renseigne sur les degrés de connexité  $CD(S_i^{1,k})$  des sous-ensembles de pixels  $S_i^{1,k}$  dont les niveaux de couleur sont définis sur les intervalles  $\left[T_{g,i}^{1,k}, T_{d,i}^{1,k}\right]$  ainsi que sur les pouvoirs discriminants  $R^{1,k}$  des composantes couleur mesurés à partir de l'image "hand".

k	1		2		3
i	1	2	1	2	1
$T_{g,i}^{1,k}$	0	120	0	67	0
$T_{d,i}^{1,k}$	119	220	66	173	179
$CD(S_i^{1,k})$	0.98	0.97	0.91	0.90	0.99
$R^{1,k}$	1.95		1.81		0.99

**Tableau 4.1 :** Degré de connexité des sous-ensembles de pixels  $S_l^{1,k}$  dont les niveaux de couleur appartiennent aux intervalles  $\left[T_{g,i}^{1,k}, T_{d,i}^{1,k}\right]$  pour l'image "hand".

Afin de pouvoir interpréter ces résultats, la figure 4.6 montre les pixels de l'image "hand" inclus dans les sous-ensembles de pixels  $S_i^{1,k}$ , k = 1,2,3. Les images des figures 4.6(a) et 4.6(b) montrent les sous-ensembles  $S_1^{1,1}$  et  $S_2^{1,1}$  de pixels associés aux deux pics significatifs détectés par l'analyse de l'histogramme de la composante rouge. En examinant visuellement ces deux figures, nous constatons que les pixels de  $S_1^{1,1}$  représentent l'anneau et le fond tandis que ceux de  $S_2^{1,1}$  représentent uniquement la main. Le pouvoir discriminant  $R^{1,1}$ , à savoir la somme des degrés de connexité de ces deux sous-ensembles, atteint 1.95 qui est une valeur très élevée.

Les images des figures 4.6(c) et 4.6(d) montrent respectivement les sous-ensembles  $S_1^{1,2}$  et  $S_2^{1,2}$  associés aux deux pics significatifs mis en évidence par l'analyse de l'histogramme de la

composante verte. Le sous-ensemble  $S_1^{1,2}$  contient une partie des pixels du fond tandis que  $S_2^{1,2}$  contient les pixels représentant la main, l'anneau et l'autre partie du fond. Les pixels de  $S_2^{1,2}$  sont relativement dispersés dans l'image. Ainsi, le pouvoir discriminant  $R^{1,2}$  atteint 1.81 et se révèle être plus faible que celui de la composante rouge. Comme un seul pic significatif est détecté par l'analyse de l'histogramme de la composante bleue (cf. image de la figure 4.6(e)), le pouvoir discriminant de cette composante atteint la valeur faible de 0.99. La composante rouge est donc la plus discriminante parmi les trois composantes R,V et B au sens de la somme des degrés de connexité. Cet exemple illustre l'intérêt d'utiliser les degrés de connexité pour mesurer le pouvoir discriminant d'une composante car il permet de mettre en évidence les composantes dont les sous-ensembles  $S_i^{j,k}$  de pixels sont les plus compacts possible dans l'image.

Le pouvoir discriminant  $R^j$  de l'espace couleur j est le pouvoir discriminant le plus élevé parmi ceux de ses trois composantes :

$$R^{j} = \max_{k=1,2,3} (R^{j,k}).$$
(4.4)



(a) Pixels de l'image "hand" appartenant à  $S_1^{1,1}$ .



(c) Pixels de l'image "hand" appartenant à  $S_1^{1,2}$ .



(b) Pixels de l'image "hand" appartenant à  $S_2^{1,1}$ .



(d) Pixels de l'image "hand" appartenant à  $S_2^{1,2}$ .



(e) Pixels de l'image "hand" appartenant à  $S_1^{1,3}$ .

**Figure 4.6 :** Pixels de l'image "hand" appartenant aux sous-ensembles  $S_i^{1,k}$  correspondant aux pics significatifs des histogrammes des composantes *R*, *V* et *B*.

### 4.5.2 Sélection de l'espace couleur

Nous proposons que l'espace couleur J le plus discriminant parmi les  $N_S$  espaces disponibles est celui dont le pouvoir discriminant est le plus élevé :

$$J = \underset{j \in [1, N_S]}{\operatorname{argmax}} (R^j). \tag{4.5}$$

Nous proposons d'illustrer la sélection de l'espace couleur le plus discriminant de l'image "hand" lors de la première itération de l'algorithme de la figure 4.1. Pour cela, le tableau 4.2 renseigne sur le pouvoir discriminant de chacun des  $N_S$  espaces couleur évalué lors de la première itération de l'algorithme.

Es	pace couleur				
	j	$R^{j}$	$R^{j,1}$	$R^{j,2}$	$R^{j,3}$
1	(R,V,B)	1.95	1.95	1.81	0.99
2	(r,v,b)	2.83	2.83	0.99	2.81
3	(X,Y,Z)	1.89	1.89	1.85	0.99
4	(x, y, z)	1.90	0.99	0.99	1.90
5	(Y, I, Q)	1.95	1.86	1.95	0.99
6	(Y,U,V)	1.96	1.86	1.96	1.95
7	(I1, I2, I3)	1.95	1.85	1.95	0.99
8	(wb, rg, by)	1.95	1.85	1.95	1.95
9	$(Y, Ch_1, Ch_2)$	1.95	1.84	1.95	1.94
10	$(L^*, a^*, b^*)$	1.94	1.75	0.99	1.94
11	$(L^*, u^*, v^*)$	2.90	1.75	2.90	1.93

**Tableau 4.2 :** Pouvoir discriminant  $R^j$  de l'espace couleur j et  $R^{j,k}$  de ses composantes mesurés à partir<br/>de l'image "hand" lors de la première itération.

Ce tableau montre que l'espace  $(L^*, u^*, v^*)$  est l'espace couleur le plus discriminant parmi les  $N_S$  espaces disponibles au sens du critère  $R^j$ . Plus particulièrement, le pouvoir discriminant de la composante  $u^*$  de cet espace est le plus élevé du tableau avec  $R^{11,2} = 2.90$ . Dans ce cas, le numéro J de l'espace le plus discriminant est fixé à 11, le numéro de l'espace  $(L^*, u^*, v^*)$ .

Le tableau 4.3 renseigne sur les pics significatifs extraits des histogrammes lissés  $h^{11,1}(x,\tau)$ (c.f. figure 4.7(a)),  $h^{11,2}(x,\tau)$  (c.f. figure 4.7(b)) et  $h^{11,3}(x,\tau)$  (c.f. figure 4.7(c)) des composantes  $L^*$ ,  $u^*$  et  $v^*$ . Ce tableau renseigne également sur les degrés de connexité  $CD(S_i^{11,k})$  des sous-ensembles de pixels  $S_i^{11,k}$  ainsi que sur les pouvoirs discriminants  $R^{11,k}$  de chacune de ces composantes.



(a) Histogramme lissé de la com-(b) Histogramme lissé de la com-(c) Histogramme lissé de la com-posante  $L^*$ .posante  $u^*$ .posante  $v^*$ .

Figure 4.7 : Histogrammes lissés de l'espace couleur  $(L^*, u^*, v^*)$  pour l'image "hand".

Le tableau 4.3 montre d'abord que parmi ces trois composantes, le nombre de pics mis en évidence est le plus élevé pour la composante  $u^*$  (il y en a 3). Par ailleurs, les degrés de connexité des 3 pics répartis sur la composante  $u^*$  sont proches de 1 et sont supérieurs ou égaux à ceux des composantes  $L^*$  et  $v^*$ . Ceci montre que les pixels formant les sous-ensembles  $S_i^{11,2}$  associés aux pics significatifs de la composante  $u^*$  sont très compacts dans l'image. Afin d'illustrer ces résultats, la figure 4.8 montre les différents sous-ensembles de pixels  $S_i^{11,2}$  de l'image "hand"

j	1		2			3	
i	1	2	1	2	3	1	2
$T_{g,i}^{11,k}$	36	4	74	107	38	70	101
$T_{d,i}^{11,k}$	64	35	106	145	73	100	122
$CD(S_i^{11,k})$	0.881	0.865	0.965	0.971	0.959	0.971	0.959
$R^{11,j}$	1.746		2.895			1.9	930

**Tableau 4.3 :** Pouvoir discriminant des sous-ensembles de pixels  $S_i^{11,k}$  dont les niveaux de couleur appartiennent aux intervalles  $[T_{g,i}^{11,k}, T_{d,i}^{11,k}]$  pour l'image "hand" codé dans le système  $(L^*, u^*, v^*)$ .



(a) Pixels de l'image "hand" ap- (b) Pixels de l'image "hand" ap- (c) Pixels de l'image "hand" appartenant à  $S_2^{11,2}$ . partenant à  $S_1^{11,2}$ . partenant à  $S_3^{11,2}$ .



correspondant aux pics délimités par les seuils détectés à partir de l'analyse de l'histogramme de la composante  $u^*$ . Sur cette figure, les couleurs des pixels appartenant à ces sous-ensembles ont été affichées dans l'espace couleur (R, V, B) pour un souci de visibilité.

En examinant visuellement cette image, nous constatons que :

- les pixels représentant le fond de l'image forment le sous-ensemble  $S_1^{11,2}$ ,
- les pixels représentant la main constituent le sous-ensemble  $S_2^{11,2}$ , les pixels représentant l'anneau forment le sous-ensemble  $S_3^{11,2}$ .

Ainsi l'analyse de la composante  $u^*$  permet de discriminer correctement les différentes régions de l'image.

L'étape 5 de l'algorithme de la figure 4.1 que nous venons de présenter sélectionne de manière non supervisée l'espace couleur le plus discriminant au sens de la somme des degrés de connexité. Les couleurs des pixels vont être codées dans cet espace pour construire une classe de pixels lors de l'étape suivante.

#### 4.6 **Construction d'une classe**

A chaque itération de l'algorithme de la figure 4.1, une classe de pixels est construite à partir de l'analyse des histogrammes lissés  $h^{J,k}(x,\tau)$  de l'espace couleur J le plus discriminant déterminé par l'étape précédente de l'algorithme. Le domaine de définition de cette classe de

pixels correspond à un parallélépipède rectangle dans l'espace couleur J. Ce parallélépipède est défini par les seuils délimitant les pics significatifs détectés pour chaque composante couleur de l'espace J. Afin de déterminer les coordonnées de ce parallélépipède dans l'espace J, nous proposons d'utiliser préférentiellement la composante couleur K la plus discriminante de l'espace J puisqu'elle permet de discriminer au mieux les pixels selon le critère utilisé. Cette composante K s'exprime par :

$$K = \operatorname*{argmax}_{k \in \{1,2,3\}} \left( R^{J,k} \right) \tag{4.6}$$

La classe de pixels est construite en deux étapes successives. La première analyse l'histogramme lissé de la composante *K*. L'intervalle de la composante *K* retenu pour construire la classe est délimité par les deux seuils  $T_{g,P^{J,K}}^{J,K}$  et  $T_{d,P^{J,K}}^{J,K}$  du pic ayant l'amplitude maximale la plus importante. La procédure ne considère alors que les pixels dont les niveaux de la composante *K* appartiennent à cet intervalle, à savoir le sous-ensemble  $S_{pJ,K}^{J,K}$  de pixels.

La seconde étape consiste à déterminer deux intervalles à partir des seuils délimitant les pics significatifs pour chacune des deux autres composantes de l'espace couleur *J*. Ces deux composantes notées  $k_1$  et  $k_2$  permettent de construire un plan couleur  $(k_1, k_2)$  de l'espace couleur *J* tel que  $k_1$  et  $k_2 \in \{1, 2, 3\}$  et  $k_1 \neq k_2 \neq K$ .

Les seuils pour ces deux composantes couleur sont sélectionnés parmi les seuils  $T_{g,i}^{J,k_1}$ ,  $T_{d,i}^{J,k_1}$ ,  $T_{g,i}^{J,k_2}$  et  $T_{d,i}^{J,k_2}$  déterminés lors de la phase de détection des pics. Ces seuils partitionnent le plan couleur  $(k_1, k_2)$  en rectangles. La procédure mesure les tailles des populations de pixels dont les points couleur projetés sur ce plan appartiennent aux différents rectangles. La projection sur ce plan couleur du domaine de définition de la classe est le rectangle avec la plus grande population.

Pour construire la classe de pixels de l'image "hand" lors de la première itération, la procédure examine l'histogramme lissé de la composante  $u^*$ . Le pic retenu, à savoir celui avec la plus grande amplitude, est délimité par les seuils 74 et 106 (cf. tableau 4.3 et figure 4.7(b)). La procédure ne considère que les pixels dont les niveaux de la composante  $u^*$  sont compris dans cet intervalle. Il reste alors à déterminer les intervalles des composantes couleur  $L^*$  et  $v^*$ pour définir le domaine de définition de la classe de pixels. La figure 4.9 montre la partition du plan couleur  $(L^*, v^*)$  à partir des seuils  $T_{g,i}^{11,1}$ ,  $T_{g,i}^{11,3}$ ,  $T_{d,i}^{11,1}$  et  $T_{d,i}^{11,3}$  définis par le tableau 4.3. Le rectangle de ce plan couleur avec la plus grande taille de population de pixels est défini par l'intervalle [4,35] pour la composante  $L^*$  et par l'intervalle [70,100] pour la composante  $v^*$ . Le domaine de définition de la classe construite à la première itération est donc délimité par l'intervalle [4,35] pour la composante  $L^*$ , [74,106] pour la composante  $u^*$  et [70,100] pour la composante  $v^*$ .



A l'issue de l'étape 6 de l'algorithme de la figure 4.1, une classe de pixels a donc été construite en analysant la distribution des points couleur dans l'espace le plus discriminant.

# 4.7 Extraction des pixels assignés à la classe construite

Les pixels dont les points couleur sont inclus dans le domaine de définition de la classe construite, sont assignés à cette classe. Ils sont alors extraits de l'image couleur afin de ne pas être pris en compte lors de l'itération suivante. Les pixels connnexes assignés à cette classe forment ainsi une partie des régions de l'image segmentée.

La figure 4.10 montre les résultats obtenus par la première itération de l'algorithme appliqué à l'image "hand". Les pixels étiquetés en blanc de l'image 4.10(a) correspondent aux pixels assignés à la première classe construite et la figure 4.10(b) montre les pixels restant à traiter pour les itérations suivantes. Les pixels de l'image assignés lors de la première itération de l'algorithme correspondent aux pixels du fond représentant les fibres textiles foncées.

L'étape 7 de l'algorithme de la figure 4.1 extrait de l'image les pixels assignés à la classe construite. Une fois l'itération terminée, il convient de déterminer si il faut construire une autre classe de pixels ou si l'algorithme doit s'arrêter.



(a) Pixels de l'image "hand" étiquetés à l'issue de (b) Pixels de l'image "hand" restant à analyser la première itération.

Figure 4.10 : Résultats de la première itération de l'algorithme de la figure 4.1 pour l'image "hand".

# 4.8 Critère d'arrêt

Nous considérons que l'algorithme doit s'arrêter lorsque un nombre suffisant de pixels de l'image ont été assignés aux différentes classes construites. Ainsi, l'algorithme continue de construire de nouvelles classes tant qu'un pourcentage p de pixels de l'image n'a pas été assigné à une classe. Ce paramètre p, ajusté par l'utilisateur, permet de moduler la finesse de la segmentation. En effet, les dernières classes extraites par l'algorithme proposé correspondent parfois à des pixels situés sur des transitions entre les régions. L'ajustement du paramètre p permet alors de ne pas les prendre en compte.

Quand l'algorithme s'arrête, les pixels qui n'ont pas été assignés à une classe lors de l'une des itérations précédentes peuvent être assignés à une des classes construites précédemment à partir d'une règle de décision.

La figure 4.11 montre les pixels de l'image "hand" assignés aux classes construites lors des itérations successives. Les pixels appartenant aux classes ainsi construites sont montrés sur les images 4.11(a), 4.11(b), 4.11(c), 4.11(g), et 4.11(h) pour respectivement les itérations 1, 2, 3, 4 et 5. Les images 4.11(d), 4.11(e), 4.11(f), 4.11(i), et 4.11(e) montrent les pixels restant à traiter pour l'itération suivante. L'image 4.11(k) montre les pixels assignés aux différentes classes. Les pixels étiquetés en noir ne sont assignés à aucune classe. Cette image montre que la première classe de pixels correspond au fond de l'image, la seconde à la main, la troisième à l'anneau ainsi qu'à une partie du fond, la quatrième à la partie de l'anneau présentant un reflet provoqué par la lumière incidente et la dernière classe à des zones de transition entre la main et le fond. Cette image montre que l'algorithme fournit des résultats de segmentation qui respectent le contenu de l'image.

Après avoir détaillé les différentes étapes ainsi que le critère d'arrêt de l'algorithme de la figure 4.1, nous montrons dans le paragraphe suivant quelques segmentations d'images de référence obtenues par cet algorithme.



(a) Pixels appartenant à la classe (b) Pixels appartenant à construite lors de la première itération.



l'issue de la première itération.



quatrième itération.



(i) Pixels qui restent à traiter à (j) Pixels qui restent à traiter à

**Résultats** 

4.9



classe construite lors de la deuxième itération.



(d) Pixels qui restent à traiter à (e) Pixels qui restent à traiter à (f) Pixels qui restent à traiter à l'issue de la deuxième itération. l'issue de la troisième itération.



(g) Pixels appartenant à la (h) Pixels appartenant à la classe construite lors de la classe construite lors de la cinquième itération.



l'issue de la quatrième itération. l'issue de la cinquième itération.

Figure 4.11 : Résultats de l'image "hand".



(k) Image segmentée.





gure 4.14(a)). Le paramètre p est fixé à 1%. On rapelle que cette valeur du paramètre p signifie qu'une nouvelle classe de pixels est extraite de l'image à segmenter tant que 99% des pixels de l'image n'ont pas été assignés à l'une des classes construites.

Les résultats obtenus par la méthode appliquée aux images "house", "peppers" et "jelly beans" seront également présentés sous la forme de tableaux (respectivement les tableaux 4.4, 4.5 et 4.6). Dans ces tableaux seront présentés pour chaque classe extraite, l'espace couleur le plus discriminant, son pouvoir discriminant  $R^J$  ainsi que le nombre de pics significatifs de la composante K la plus discriminante. Enfin, les pixels assignés à chacune des classes construites seront étiquetés avec des fausses couleurs spécifiques aux classes.

#### 4.9.1 Image "house"

L'image originale "house" présente plusieurs difficultés en vue d'une segmentation à cause notamment des effets d'ombre sous le toit de la maison. De plus, les couleurs des pixels représentant le mur et le toit sont hétérogènes (cf. image de la figure 4.12(a)). Une image correctement segmentée doit par conséquent permettre de retrouver ces différentes régions. L'image segmentée retournée par la méthode proposée (cf. image de la figure 4.12(b)) est constituée de huit classes correspondant au ciel, aux murs, aux parties ombrées du mur, au toit, aux parties ombrées des gouttières, aux gouttières, à la fenêtre principale et à des pixels relativement dispersés dans l'image.

Le tableau 4.4 montre que ce sont les espaces couleur (Y,U,V), (R,V,B), (bw,rg,by), (I1,I2,I3) et (r,v,b) qui ont été sélectionnés à l'une des itérations. Cet exemple met ainsi en évidence l'intérêt de déterminer un espace couleur approprié pour extraire chaque classe de pixels.

Le pouvoir discriminant de l'espace sélectionné à l'itération 3 est égal à 3.48, valeur supérieure à celles des espaces sélectionnés aux deux premières itérations. Ceci s'explique par la mise en évidence de nouveaux pics significatifs lors de la troisième itération.

Par ailleurs, nous constatons que le pouvoir discriminant de l'espace sélectionné lors de l'itération 4 (2.52) est légèrement plus faible à celui sélectionné lors de l'itération 2 (2.54) alors que le nombre de pics détéctés y est plus important. Cela signifie que les pixels appartenant aux pics significatifs sont de moins en moins connexes dans l'image au fur et à mesure des itérations. En effet, puisque de moins en moins de pixels de l'image sont considérés au fil des itérations, les pixels ne constituant pas de régions significatives dans l'image interviennent de manière de plus en plus conséquente dans le calcul du pouvoir discriminant.

De plus, nous pouvons remarquer sur cet exemple que la classe construite lors de l'itération 7 correspond à des pixels dispersés dans l'image tandis que la dernière classe construite à l'itération 8, correspond à une petite région de l'image. Par conséquent, il serait intéressant d'établir un critère de sélection des classes basé sur les degrés de connexité plutôt que sur leurs populations afin d'inverser l'ordre de construction des classes.

Image				
Itération i	1	2	3	4
J	(Y,U,V)	(Y,U,V)	(R,V,B)	(R,V,B)
$N_P^{J,K}$	4	3	5	4
$R^{J}$	3.42	2.54	3.48	2.52
Image				
Itération i	5	<mark>6</mark>	7	8
J	(bw, rg, by)	$(I_1,I_2,I_3)$	(r,v,b)	(Y,U,V)
$N_P^{J,K}$	3	3	2	2
$R^{J}$	1.75	1.94	1.20	1.49

**Tableau 4.4 :** Espace couleur *J* sélectionné à chaque itération, pouvoir discriminant  $R^{I}$  et nombre  $N_{P}^{J,K}$  de pics de la composante *K* la plus discriminante pour chaque itération de la procédure appliquée à l'image "house".



(a) Image originale "house".



(b) Segmentation de l'image "house" par la méthode proposée.

Figure 4.12 : Segmentation de l'image "house" avec la méthode proposée.

# 4.9.2 Image "peppers"

L'image originale "peppers" (cf. image de la figure 4.13(a)) est constituée de poivrons et de piments ayant des couleurs allant du vert au rouge en passant par un vert-jaune. La diffi-

culté principale de cette image provient des reflets présents sur les poivrons donnant naissance à des nuages de points mélangés dans l'espace (R, V, B). La méthode proposée construit quatre classes de pixels correspondant aux poivrons rouges, au poivrons verts foncés, aux poivrons verts-jaunes, aux reflets de l'éclairage incident sur ces légumes et construit deux classes de pixels représentant le fond (cf. image de la figure 4.13(b)).

Pour segmenter cette image, les espaces couleur (Y, U, V), (R, V, B), (X, Y, Z), (bw, rg, by), (I1, I2, I3) sont sélectionnés à l'une des itérations de la méthode (cf. tableau 4.5). Cet exemple montre à nouveau l'intérêt de déterminer l'espace couleur le plus discriminant pour construire chacune des six classes de pixels.

Le tableau 4.5 montre que le pouvoir discriminant de l'espace sélectionné à l'itération 2 est égal à 1.94 alors qu'il atteint la valeur 2.69 lors de la troisième itération. L'augmentation du pouvoir discriminant s'explique par l'augmentation entre les deux itérations du nombre de pics mis en évidence sur les composantes les plus discriminantes. En effet, on dénombre 3 pics significatifs pour la troisième itération alors que l'on en dénombre seulement 2 à l'itération précédente. La valeur élevée du pouvoir discriminant de l'espace sélectionné lors de la troisième itération prouve que les degrés de connexité des pics atteignent des valeurs élevées et donc que les pics ainsi révélés correspondent à des pixels connexes dans l'image.

Image				
Itération i	1	2	3	4
J	(Y,U,V)	(R,V,B)	(X,Y,Z)	(bw, rg, by)
$N_P^{J,K}$	3	2	3	3
$R^{J}$	2.79	1.94	2.69	2.60
Image				
Itération i	5	<mark>6</mark>		
J	$(I_1, I_2, I_3)$	(R,V,B)		
$N_P^{J,K}$	2	2		
$R^{J)}$	1.84	1.42		

**Tableau 4.5 :** Espace couleur *J* sélectionné à chaque itération, pouvoir discriminant R' et nombre  $N_P^{J,K}$  de pics de la composante *K* la plus discriminante pour chaque itération de la procédure appliquée à l'image "peppers".



(a) Image originale "peppers"



(b) Segmentation de l'image "peppers" par la méthode proposée.

Figure 4.13 : Segmentation de l'image "peppers" avec la méthode proposée.

### 4.9.3 Image "jelly beans"

L'image originale "jelly beans" est composée d'un fond et de bonbons de couleurs jaune, rouge, verte et marron (cf. image de la figure 4.14(a)). La segmentation de cette image présente plusieurs difficultés à cause de l'hétérogénéité de l'intensité des pixels représentant le fond ainsi que des ombres présentes entre les bonbons et des reflets de la lumière incidente sur ces derniers. Une image correctement segmentée doit isoler le fond des bonbons et les bonbons de couleurs différentes doivent être correctement séparés. L'image segmentée retournée par la méthode proposée est constituée de 7 classes correspondant au fond, aux bonbons marrons, jaunes, rouges, verts, aux ombres entre les bonbons et à un ensemble de pixels dispersés dans l'image.

Le tableau 4.6 montre que pour segmenter cette image, les espaces couleur (X, Y, Z),  $(Y, Ch_1, Ch_2)$  et (bw, rg, by) sont sélectionnés pour construire les sept classes de pixels.

Ce tableau montre que l'extraction des classes de pixels par itérations successives est à nouveau justifiée dès la deuxième itération. En effet, l'extraction des pixels appartenant au fond de l'image permet de révéler de nouveaux pics qui sont identifiés lors des itérations suivantes pour extraire les différentes classes de bonbons. Nous remarquons que le pouvoir discriminant de l'espace sélectionné passe de 1.95 à 3.38 entre la première et la seconde itération. Les pics révélés à la deuxième itération correspondent effectivement à des régions représentant des bonbons de l'image .

De plus, nous pouvons remarquer sur cet exemple que la dernière classe construite correspond à des pixels dispersés dans l'image. Le pouvoir discriminant de l'espace sélectionné lors de cette dernière itération atteint une valeur faible, 0.38. Ainsi, un critère d'arrêt basé sur les degrés de connexité plutôt que sur le nombre de pixels à traiter permettrait d'arrêter l'algorithme itératif et par conséquent d'éviter d'extraire des classes non significatives.

Après avoir détaillé les résultats obtenus par notre algorithme appliqué aux images "house", "peppers" et "jelly beans", nous regroupons les différentes conclusions obtenues à partir de l'analyse de ces résultats dans la synthèse suivante.

Image				
Itération i	1	2	3	4
J	(X,Y,Z)	$(Y, Ch_1, Ch_2)$	(bw, rg, by)	(bw, rg, by)
$N_P^{J,K}$	2	4	3	3
$\hat{R}^J$	1.95	3.38	2.71	2.72
Image				
Itération i	5	<mark>6</mark>	7	
J(I)	(bw, rg, by)	(bw, rg, by)	(bw, rg, by)	
$N_P^{J,K}$	2	2	1	
$\hat{R}^J$	1.86	1.36	0.38	

**Tableau 4.6 :** Espace couleur *J* sélectionné à chaque itération, pouvoir discriminant  $R^{I}$  et nombre  $N_{P}^{J,K}$  de pics de la composante *K* la plus discriminante pour chaque itération de la procédure appliquée à l'image "jelly beans".



(a) Image originale "jelly beans".



(b) Segmentation de l'image "jelly beans" par la méthode proposée.


#### 4.9.4 Synthèse des résultats

La principale originalité de notre méthode de segmentation est la sélection de l'espace couleur le plus discriminant pour la construction de chaque classe de pixels. Cet algorithme itératif opère dans un contexte non supervisé. Les tableaux 4.4, 4.5 et 4.6 résultant de la segmentation respective des images "house", "peppers" et "jelly beans" montrent qu'il n'existe pas un unique espace discriminant selon notre critère basé sur la somme des degrés de connexité. Ceci justifie expérimentalement la sélection de l'espace couleur le plus discriminant pour chaque itération de l'algorithme.

Nous remarquons également que l'aspect itératif de l'algorithme permet de construire, lors d'une itération donnée, des classes qui n'étaient pas détectables lors des itérations précédentes. Ceci est mis en évidence par l'augmentation du nombre de pics significatifs  $N_P^{J,K}$  entre deux itérations successives. Par exemple, le nombre de pics détectés passe de 3 à 5 entre la deuxième et la troisième itération de notre algorithme appliqué à l'image "house".

Les différents résultats obtenus montrent que notre algorithme permet dans sa version actuelle d'obtenir des résultats de segmentation encourageants. Nous n'avons pas pris le temps d'approfondir cette démarche et notamment de comparer les résultats obtenus par notre approche avec ceux obtenus par d'autres méthodes. Il nous faudra donc mesurer de manière quantitative les résultats obtenus. Deux solutions existent. D'une part, il faut appliquer notre méthode sur des images de synthèse dont on connaît la vérité-terrain [Fon01]. D'autre part, il faut mesurer avec des critères de qualité de segmentation, tels que ceux proposés par Borsotti [BCS98], les résultats obtenus par notre approche et comparer les mesures obtenues avec celles obtenues par d'autres méthodes de segmentation.

## 4.10 Conclusion

Le choix de l'espace couleur dans lequel sont codées les couleurs des pixels traités par un algorithme de segmentation est un problème actuellement ouvert. Dans ce chapitre, nous apportons notre contribution en proposant une méthode de segmentation par multi-seuillage. Sa principale originalité est la sélection non supervisée de l'espace couleur le plus discriminant pour construire des classes de pixels en présence dans l'image.

L'un des principaux défauts des méthodes de multi-seuillage par analyse d'histogrammes est qu'elles ne prennent pas en compte l'arrangement spatial des pixels dans l'image. Aussi, l'algorithme que nous proposons exploite une mesure sur l'arrangement spatial des pixels pour mesurer le pouvoir discriminant de chaque espace couleur. D'autre part et contrairement aux approches de Ohlander et de Vandenbroucke, la construction des classes s'effectue dans l'espace couleur 3D et non pas seulement sur une seule composante ou un espace hybride sans réalité physique, physiologique ou psychologique.

La version actuelle de notre algorithme fournit des résultats prometteurs qui peuvent être améliorés par la modification des points suivants.

La qualité de lissage des histogrammes est un point essentiel pour obtenir des images correctement segmentées. La méthode de Lin utilisée actuellement lisse de manière adaptative les histogrammes des composantes lors de l'étape numéro 2. Il s'est avéré que les histogrammes résultants sont sur-lissés, provoquant parfois le regroupement de deux pics significatifs voisins le long d'une composante. Il nous semble opportun d'intégrer le spatial scale space filtering décrit au chapitre précédent à notre algorithme. Ceci permettra sans doute d'améliorer la qualité de lissage des histogrammes et donc la qualité de détection des seuils délimitant les pics significatifs.

Le critère d'arrêt actuellement utilisé est basé sur la taille de la population des pixels assignés aux classes construites lors des itérations précédentes. Il serait intéressant que ce critère prenne en considération le pouvoir discriminant de l'espace sélectionné lors de l'itération courante. En effet, lorsque l'espace couleur sélectionné est faiblement discriminant, il est probablement opportun d'arrêter l'algorithme afin de ne pas construire des classes de pixels non significatives dans l'image.

Enfin, le pouvoir discriminant d'un espace couleur tri-dimensionnel est mesuré en considérant chacune de ses composantes de manière marginale. Il serait intéressant de généraliser cette mesure en considérant les degrés de connexité de sous-ensembles de pixels définis par des parallélépipèdes rectangles de l'espace couleur, ces parallélépipèdes étant délimités par les seuils détectés le long des composantes couleur.

Néanmoins, cette approche a principalement été développée pour répondre à un problème industriel de contrôle qualité. Nous verrons dans le chapitre suivant que notre approche a été intégrée dans un système industriel de contrôle d'aspect et qu'elle fournit d'excellents résultats en termes de détection de défauts apparaissant sur des motifs colorés de verres sérigraphiés.

# **Chapitre 5**

# Application à la détection de défauts sur verres serigraphiés

Le contrôle qualité de la production fait partie actuellement des enjeux stratégiques des entreprises car il permet d'assurer au client le niveau d'exigence requis sur la qualité du produit manufacturé. De nombreux contrôles sont effectués par les entreprises comme par exemple le contrôle dimensionnel ou encore le contrôle de présence-absence. Ces contrôles sont couramment effectués par des systèmes de vision dans lesquels des algorithmes performants sont implantés. D'autres types de contrôles comme le contrôle d'aspect à partir d'images couleur sont plus délicats à effectuer et très peu d'outils sont disponibles sur le marché des logiciels de traitements d'images dédiés aux industriels. C'est dans ce domaine que se situent les travaux présentés dans ce chapitre.

L'entreprise Arc International, leader mondial des arts de la table produit des verres ornés de motifs imprimés par sérigraphie. La sérigraphie sur verre est une application d'émail de teintes différentes sur un support en verre creux et cylindrique. Le principe consiste à faire tourner le support en verre neutre sous 1 à 8 écrans de soie qui déposent chacun un émail de couleur. Cuits à environ 650° C, les émaux se vitrifient sur la première peau du verre. Le décor prend alors son éclat et devient indélébile. Cependant, plusieurs types de défauts liés à cette technologie peuvent apparaitre. Les verres présentant de tels défauts ne peuvent être vendus au client et doivent être sortis du processus de fabrication.

Actuellement un contrôle de la qualité des décors est effectué visuellement par des opérateurs humains. Le contrôle qualité par une inspection visuelle d'articles tels que les verres est une tâche éprouvante pour l'opérateur qui ne peut garantir la répétabilité du contrôle. De plus, les résultats de ce contrôle peuvent être très différents car ils dépendent de nombreux paramètres tels que la fatigue, l'état de stress, le rythme de production, la complexité des contrôles à effectuer... Enfin, le travail effectué par l'opérateur est un contrôle qualité qui n'apporte aucune valeur ajoutée au produit. Pour ces raisons, il est préférable d'automatiser cette phase de contrôle.

Dans ce but, nous proposons de mettre en place un système de vision industrielle permettant le contrôle automatique des décors des verres imprimés par sérigraphie. Ce travail consiste à mettre au point le poste d'acquisition des images couleur puis d'appliquer différents traitements sur l'image couleur afin de détecter les différents types de défauts.

La présentation du système de vision proposé dans ce chapitre est divisée en quatre parties. Dans la première partie de ce chapitre, nous présentons le poste d'acquisition spécifiquement développé pour le contrôle qualité. Ce poste d'acquisition permet d'acquérir les images des verres imprimés par sérigraphie pour lesquels nous présenterons les différents types de défauts dans la deuxième partie. La troisième partie présentera alors les différents traitements proposés afin de les détecter. Enfin, la quatrième partie présentera quelques résultats obtenus.

# 5.1 Présentation du poste d'acquisition

Dans ce paragraphe, nous détaillons les différents éléments constituant le poste d'acquisition développé pour acquérir les images des verres imprimés par sérigraphie. La figure 5.1 montre deux photographies du poste d'acquisition qui mettent en avant les différents organes mécaniques utilisés. L'image de la figure 5.1(a) montre une vue de face du poste d'acquisition sur laquelle sont repérés :

- le système de préhension reliant le verre à l'arbre moteur,
- l'arbre moteur reliant le système de préhension au moteur,
- le codeur incrémental qui est fixé sur l'arbre moteur.

L'image de la figure 5.1(b) montre une vue de coté du poste d'acquisition sur laquelle sont repérés :

- la fibre optique permettant d'éclairer le verre,
- la caméra linéaire couleur,
- un filtre infra-rouge placé devant l'objectif de la caméra.



(a) Vue de face du poste d'acquisition.



(b) Vue de côté du poste d'acquisition.

Figure 5.1 : Poste d'acquisition.

Nous pouvons voir sur ces photographies que le verre inspecté est maintenu verticalement par un système de préhension. Ce système de préhension est relié à un moteur via l'axe moteur qui assure la rotation du verre autour de son axe principal d'inertie à la vitesse de 1 tour par seconde. L'acquisition d'images-ligne successives permet alors d'obtenir une représentation bidimensionnelle de la surface du verre grâce à sa rotation. La figure 5.2 montre le principe de fonctionnement du montage mécanique utilisé pour notre étude tel que nous venons de l'énoncer.



Figure 5.2 : Principe de fonctionnement du poste d'acquisition.

Après avoir présenté le système mécanique du poste d'acquisition, les paragraphes suivants ont pour objectifs de présenter les différents éléments de la chaîne d'acquisition. Ainsi, le paragraphe suivant présente la caméra linéaire utilisée. Le second paragraphe présente le système d'éclairage et le dernier paragraphe la focale choisie.

#### 5.1.1 La caméra linéaire couleur

La caméra utilisée est une caméra linéaire couleur Basler L301 - kc de fréquence maximale d'acquisition  $F_{max} = 9.2$ kHz équipée d'un capteur Kodak KLI - 2113 [BAS03]. Ce capteur est équipé de trois lignes de capteur CCD de  $N_t = 2098$  photo-éléments chacune. Les récepteurs photo-sensibles sont des surfaces carrées de 14  $\mu$ m ×14  $\mu$ m de côté et chaque ligne de capteur est distante de 112  $\mu$ m. Les courbes de sensibilités spectrales des trois types de capteurs R, Vet B sont représentées sur la figure 5.3.



**Figure 5.3 :** Courbes de sensibilité spectrale du capteur linéaire couleur Kodak KLI-2113 équipant la caméra Basler L301 - kc (données constructeur).

L'axe des ordonnées de cette figure est le rendement quantique du capteur. Le rendement quantique d'un capteur est définie par le rapport  $\frac{\text{électrons produits}}{\text{photons incidents}}$  et correspond au taux de conversion entre les photons incidents et les électrons produits à une longueur d'onde donnée. La figure 5.3 montre que les capteurs sont sensibles à des longueurs d'ondes extérieures au domaine du visible (380, 750 nm) et particulièrement aux infra-rouges. C'est pourquoi nous utilisons un filtre IR afin que les rayons infra-rouges ne sensibilisent pas les capteurs et ne viennent ajouter de l'information ne correspondant pas aux motifs couleur du verre perçu par l'oeil humain.

Le cahier des charges impose d'inspecter des objets d'une longueur h = 200 mm avec une durée d'acquisition de l'image de la surface du verre à inspecter inférieure à la seconde. La résolution transversale est par conséquent fixée à  $R_t = \frac{h}{N_t} = \frac{200}{2098} = 0.095$  mm/pixel. La fréquence d'acquisition est réglée de telle sorte à obtenir une précision longitudinale égale à la précision transversale.

Après avoir présenté la caméra linéaire couleur, le paragraphe suivant détaille le système d'éclairage.

#### 5.1.2 Le système d'éclairage

Pour cette application, nous avons utilisé une source de lumière froide de la marque EU-ROSEP 5050 Spotlight [EUR03]. Cette source de lumière est un éclairage de type continu avec une ampoule à halogénure métallique. Cette source de lumière émet dans l'ensemble du spectre du visible ce qui permet de sensibiliser toutes les longueurs d'ondes du spectre et caractériser ainsi parfaitement la couleur du matériau. D'autre part la source de lumière émet très peu dans l'infra-rouge comme le montre sa courbe de répartition spectrale de la figure 5.4. Malgré cela, il est tout de même nécessaire de placer un filtre IR tant la sensibilité du capteur rouge y est importante par rapport aux capteurs vert et du bleu comme le montre la courbe de sensibilité spectrale de la caméra (c.f. figure 5.3).



**Figure 5.4 :** Représentation spectrale relative d'énergie de la lampe à halogénure-métallique utilisée (courbe constructeur).

Une fibre optique est placée à la sortie de cette source lumineuse pour mettre en forme la lumière. Cette fibre optique permet de répartir la lumière le long d'une ligne projetée sur la surface du verre à inspecter via un diffuseur d'une hauteur de 23 cm dans l'axe optique de la caméra. La géométrie d'éclairage utilisée est ainsi du type bright-field (c.f. paragraphe 1.2.2.2). Selon l'ajustement du diffuseur à la sortie de la fibre optique, la lumière projetée sur le verre est focalisée sur une bande de largeur plus ou moins importante et par conséquent avec une intensité lumineuse plus ou moins forte. Ce diffuseur qui est de plus grande taille que le verre à inspecter permet une répartition homogène de la lumière le long de la ligne de visée de la caméra en évitant une diminution de l'intensité aux extrémités de la bande éclairée. La lumière ainsi focalisée sur le verre doit satisfaire aux trois exigences suivantes :

- La ligne de visée de la caméra ne doit pas se situer dans la direction de la réflexion spéculaire sous peine que le capteur reçoive directement les rayons lumineux provenant de la source de lumière. Dans ce cas, aucune information sur la couleur des motifs du verre ne peut être restituée. Cependant, la ligne de visée de la caméra doit être proche de la direction de la réflexion spéculaire afin que le capteur perçoive les rayons réfléchis de manière diffuse avec un maximum d'intensité.
- La surface de la bande du verre éclairée ne doit pas être trop large. En effet l'intensité de la lumière incidente dépend de la surface éclairée. Quand cette surface est trop importante, la puissance lumineuse est trop faible, ce qui a pour conséquence de diminuer le rapport signal sur bruit. Ceci se traduit par l'acquisition d'images-ligne bruitées.
- A cause de contraintes technologiques liées à leur fabrication, les verres ne sont ni parfaitement concentriques, ni parfaitement coniques. La surface du verre éclairée par la fibre

optique ne doit pas par conséquent être trop fine sous peine que la ligne de visée de la caméra ne soit plus localisée dans cette surface éclairée si le verre présente des défauts géométriques.

Ce paragraphe détaille le système d'éclairage utilisé pour cette application industrielle. Un point important qui doit être pris en compte sont les défauts géométriques des verres à analyser. Ces défauts géométriques doivent aussi être pris en considération pour le calcul de la focale de l'objectif équipant la caméra.

#### 5.1.3 Le choix de la focale

Pour déterminer la focale de l'objectif qui équipe la caméra, nous devons satisfaire à deux exigences contradictoires. En effet, il faut d'une part que la caméra soit la plus éloignée possible de l'article inspecté afin de minimiser les problèmes d'acquisition causés par les défauts géométriques des verres. En effet, les défauts géométriques sont d'autant plus visibles que la caméra est proche du produit à inspecter. D'autre part il faut que la caméra soit suffisamment proche de l'objet pour que la puissance des stimuli reçus soit la plus importante possible. En effet, l'éclairement d'une surface est inversement proportionnel au carré de la distance entre cette surface et la source lumineuse [Gau06]. Aussi, afin d'aboutir à un compromis face à ces deux exigences contradictoires, nous avons décidé d'utiliser un objectif avec une focale f de 75 mm. La caméra est alors située à une distance d de 511 mm comme le montre l'équation 5.1, où c correspond à la longueur du capteur de la caméra.

$$d = \frac{f \times h}{c} = \frac{75 \times 200}{2098 \times 14.10^{-3}} \approx 511 \text{mm}$$
(5.1)

Après avoir présenté les différents éléments constituant le poste d'acquisition, nous présentons les défauts provoqués par un dysfonctionnement du dispositif d'impression par sérigraphie des verres.

## 5.2 Présentation des défauts

Dans ce paragraphe nous présentons les quatre types de défauts que le système de contrôle d'aspect doit détecter :

- Les défauts appelés "tâchés" sont des tâches de couleur sur les verres. Les "tâchés" ont deux origines :
  - 1. une goutte de couleur est tombée sur le verre lors de l'étape de sérigraphie,
  - 2. deux verres qui n'étaient pas encore sec se sont touchés.

La figure 5.5 montre de tels défauts. L'image de la figure 5.5(a) est une image étalon sur laquelle il n'a pas de défaut. Les images des figures 5.5(b) et 5.5(c) montrent respectivement l'image d'un verre avec un défaut de "tâché" lorsqu'une goutte de couleur est tombée et lorsque deux verres qui n'étaient pas encore secs se sont touchés.



 (a) Image d'un verre sans défaut.
(b) Image d'un verre avec un dé- (c) Image d'un verre avec un défaut de "tâché".
faut de "tâché".

Figure 5.5 : Illustration d'un verre avec un défaut de "tâché".

Les défauts appelés "manques" sont repérables par l'absence de tout ou d'une partie d'un écran lors de la sérigraphie du verre. La figure 5.6 montre un verre exempt de défaut sur l'image 5.6(a) et un verre avec un défaut de "manque" sur l'image 5.6(b).



(a) Image d'un verre sans défaut.

(b) Image d'un verre avec un défaut de "manque".

Figure 5.6 : Illustration d'un verre avec un défaut de "manque".

Les défauts dits de "décalage" sont provoqués par le décalage de deux écrans différents lors de la sérigraphie du verre. La figure 5.7 montre un verre exempt de défaut sur l'image 5.7(a) et un verre avec un défaut de "décalage" sur l'image 5.7(b).



(a) Image d'un verre sans défaut.

(b) Image d'un verre avec un défaut de "décalage".

Figure 5.7 : Illustration d'un verre avec un défaut de "décalage".

Les défauts de "nuance" sont provoqués par une couleur d'écran non homogène sur l'ensemble de la surface imprimée par l'écran. Ils peuvent par exemple provenir d'un manque d'encre lors de l'impression du verre. La figure 5.8 montre un verre exempt de défaut sur l'image 5.8(a) et un verre avec un défaut de "décalage" sur l'image 5.8(b).



(a) Image d'un verre sans défaut.

(b) Image d'un verre avec un défaut de "nuance".

Figure 5.8 : Illustration d'un verre avec un défaut de "nuance".

# 5.3 Contrôle d'aspect par analyse d'images couleur

La fabrication d'un verre est une tâche complexe. Par exemple, la géométrie du verre dépend de la manière dont tombe la goutte de verre en fusion dans le moule. Cette complexité de fabrication empêche d'appliquer une simple soustraction d'images entre une image de référence et l'image du verre à inspecter sous peine de faire apparaître de nombreux effets de bord. Le tableau 5.3 illustre le résultat d'une telle soustraction. Les images 5.0(a) et 5.0(b) sont les images opérandes et l'image 5.0(c) est l'image résultante de la soustraction des deux images opérandes. Bien que les images opérandes 5.0(a) et 5.0(b) représentent deux verres différents ayant le même décor, celles-ci contiennent des zones identiques situées aux mêmes coordonnées spatiales puisque les images des verres dont elles sont extraites ont été préalablement recalées suivant l'axe des ordonnées qui représente aussi la direction longitudinale.



Tableau 5.1 : Soustraction de deux images.

Bien qu'un défaut (une petite tâche blanche) présent dans l'image opérande 5.0(a) soit présent dans l'image résultante 5.0(c), il semble difficile de le détecter tant le contenu de cette image résultante est encore riche. Cet exemple montre ainsi qu'il est impossible de soustraire une image de référence à l'image du verre à inspecter à cause des déformations géométriques des verres. Cependant, nous avons remarqué expérimentalement qu'analyser des parties d'images de tailles réduites que nous appellerons *régions d'intérêt*<sup>1</sup> permet de minimiser ces problèmes d'effets de bord lorsque celles-ci sont recherchées dans les images par un outil de corrélation normalisée.

En outre, la cadence de production élevée des verres impose un temps de traitement court. Pour cette raison, nous avons décidé d'utiliser un algorithme opérant en deux étapes successives. La première est une phase d'apprentissage supervisé nécessitant un nombre de calculs important par rapport au temps de traitement disponible pour le contrôle du verre. Pour cette raison, la phase d'apprentissage est effectuée hors-ligne. La seconde étape exploite les résultats acquis lors de la phase d'apprentissage pour décider si les motifs imprimés sur le verre à inspecter ne possèdent aucun des différents défauts cités dans le paragraphe 5.2. Cette procédure qui opère sur la ligne de production est effectuée en-ligne.

Afin de contrôler les motifs des verres, nous définissons dans un premier temps une ROI à partir d'un verre de référence exempt de défauts lors de la phase d'apprentissage. Cette ROI que nous appelons *ROI de référence* est définie interactivement par l'opérateur. Ce dernier doit également spécifier le nombre  $N_C$  de classes de pixels qui la constituent. Ces  $N_C$  classes amènent à deux approches différentes pour détecter les défauts suivant que la ROI de référence ne possède qu'une seule ( $N_C = 1$ ) ou plusieurs ( $N_C > 1$ ) classes. En effet, dans le cas où  $N_C = 1$  aucune information supplémentaire n'est nécessaire par la méthode proposée tandis que pour le cas où  $N_C > 1$  une approche similaire à celle présentée dans le chapitre précédent est effectuée. Ainsi, pour le cas  $N_C > 1$ , les espaces couleur les plus discriminants ainsi que les seuils nécessaires à la construction de chaque classe sont déterminés lors de la phase d'apprentissage et seront exploités pour la procédure en-ligne.

La première étape de la procédure en-ligne consiste à rechercher dans le verre à inspecter la ROI de référence définie interactivement par l'opérateur lors de la phase d'apprentissage.

<sup>&</sup>lt;sup>1</sup>Region Of Interest (ROI) en anglais.

Cette recherche de la ROI de référence dans le verre à inspecter est effectuée en appliquant un algorithme de corrélation normalisée. Si la ROI de référence est retrouvée dans le verre à inspecter, deux cas peuvent se produire suivant que le nombre  $N_C$  de classes est égal ou supérieur à 1. Si  $N_C = 1$ , un gradient couleur est calculé dans la ROI trouvée pour décider si la pièce possède un défaut. Si  $N_C > 1$ , des classes de pixels sont construites à partir de la ROI trouvée avec les espaces couleur les plus discriminants et les seuils déterminés lors de la phase d'apprentissage. Les classes de pixels ainsi construites sont comparées avec celles construites par l'analyse de la ROI du verre de référence pour décider si le verre à inspecter possède des défauts ou non. Ce sont donc les images segmentées qui sont comparées.

Les paragraphes suivants ont pour objectif de détailler cette méthode de contrôle d'aspect. Ainsi, le premier paragraphe détaille la procédure hors-ligne tandis que le deuxième détaille la procédure en-ligne.

#### 5.3.1 Procédure hors-ligne

La figure 5.9 montre les différentes étapes de la procédure hors-ligne qui est issue de la méthode présentée au chapitre 4. Seule la mesure du pouvoir discriminant de chaque espace couleur et le critère d'arrêt de la procédure itérative ont été adaptés à la problématique indus-trielle.



Figure 5.9 : Algorithme de la procédure hors-ligne.

Chaque étape de l'algorithme de la figure 5.9 sera présentée une à une dans les paragraphes suivants.

#### 5.3.1.1 Acquisition d'une image de référence

La première étape de la phase d'apprentissage consiste à acquérir l'image matricielle de la surface d'un verre de référence exempt de défauts. Cette image matricielle est la juxtaposition des différentes images-ligne acquises successivement. L'image ainsi acquise est utilisée pour que l'opérateur sélectionne de manière interactive la ROI de référence durant l'étape n°2 de l'algorithme de la figure 5.9.

#### 5.3.1.2 Sélection de la ROI de référence

Comme nous l'avons énoncé précédemment, la soustraction de l'image d'un verre de référence avec l'image d'un verre à inspecter ne fournit pas des résultats exploitables à cause des déformations des verres. Cependant, nous avons remarqué expérimentalement qu'il est possible de comparer une ROI de référence provenant de l'image de référence à celle de l'image d'un verre à inspecter sans que les déformations ne soient trop contraignantes. Aussi, cette étape de l'algorithme consiste à positionner une ROI sur l'image de référence pour déterminer la zone du verre que l'on souhaite analyser.

En outre, comme l'indique l'algorithme de la figure 5.9, deux voies sont possibles suivant que la ROI possède une ou plusieurs classes à construire. Ainsi, pour le cas où la ROI est constituée d'une seule classe, il est impossible de la retrouver dans le verre à inspecter avec les outils classiques de corrélation normalisée. Pour cette raison, l'opérateur doit placer sur l'image de référence deux ROI différentes. La première, appelée *ROI maître* doit impérativement être positionnée sur une partie du motif possédant suffisamment d'informations pour être retrouvée dans les images considérées. Cette première ROI sert par conséquent à se positionner à un endroit spécifique du verre à analyser pour s'affranchir des déformations géométriques. La seconde ROI, appelée *ROI esclave* est celle dans laquelle on examinera la présence éventuelle des défauts. La position de cette ROI fixée par l'opérateur, doit être la plus près possible de la ROI maître afin de minimiser les décalages dus aux déformations géométriques du verre. La ROI esclave du verre de référence correspond à la ROI de référence qui sera analysée. Celle-ci est repérée par rapport à la ROI maître par un décalage  $\Delta x$  suivant l'axe des ordonnées et d'un décalage  $\Delta y$  suivant l'axe des abscisses dans l'image.

La figure 5.10 illustre le positionnement d'une ROI maître et d'une ROI esclave pour un verre de référence "Fanta" sur l'image 5.10(a). La ROI maître est délimitée par un rectangle vert tandis que la ROI esclave est délimitée par un rectangle bleu. Cette ROI esclave (c.f. 5.10(b)) du verre de référence qui correspond à la ROI de référence, nous servira pour illustrer les différentes étapes effectuée par la méthode. Cette figure qui montre une ROI esclave de dimension  $147 \times 132$  permet aussi de préciser sur les dimensions des ROI que nous préconisons.



(a) Positionnement des ROI maître et esclave.

(b) ROI esclave.



Bien évidemment, si l'utilisateur ne sélectionne pas de ROI, l'algorithme s'arrête puisqu'aucune zone du verre n'est à analysée. Sinon l'étape n°3 de l'algorithme de la figure 5.9 est exécutée.

#### 5.3.1.3 Saisie du nombre de classes

Cette étape nécessite l'intervention de l'opérateur. Celui-ci doit saisir le nombre  $N_C$  de classes différentes présentes dans la ROI à analyser. Si le nombre de classes  $N_C$  est égal à 1, alors la phase d'apprentissage s'arrête. Sinon, l'algorithme va en étape 4 pour déterminer l'espace couleur le plus discriminant pour construire chaque classe de pixels.

Par exemple, pour la ROI de référence de la figure 5.10(b), on dénombre  $N_C = 5$  classes de pixels différentes correspondant aux couleurs suivantes :

- 1. orange foncé,
- 2. orange clair,
- 3. blanc,
- 4. gris,
- 5. bleu foncé.

Une fois la ROI sélectionnée et le nombre de classes spécifié (avec  $N_C > 1$ ), on recherche l'espace couleur le plus discriminant lors de l'étape n°4 pour construire chaque classe de pixels de l'algorithme.

#### 5.3.1.4 Calcul des histogrammes 1D et des histogrammes lissés

Pour déterminer l'espace couleur le plus discriminant, on analyse le pouvoir discriminant de chaque composante couleur à partir de l'analyse de leurs histogrammes. Cette étape n°4 consiste à calculer les histogrammes  $h^{j,k}(x)$  de la  $k^{\text{ème}}$  composante couleur (avec  $k \in \{1,2,3\}$ )

des  $N_S$  espaces couleur répertoriés dans le paragraphe 4.2 (avec  $j \in \{1, ..., N_S\}$  et  $N_S = 11$ ). La procédure de lissage proposée par Lin est appliquée afin de révéler les pics significatifs des histogrammes lissés  $h^{j,k}(x,\tau)$  [LWY96]. Le calcul des histogrammes  $h^{j,k}(x)$  et  $h^{j,k}(x,\tau)$ sont expliqués plus en détail dans les paragraphes 3.1 et 3.5.1 respectivement. A partir des histogrammes  $h^{j,k}(x,\tau)$  ainsi calculés, l'espace couleur le plus discriminant est sélectionné lors de l'étape n°6 de l'algorithme de la figure 5.9.

#### 5.3.1.5 Sélection de l'espace couleur le plus discriminant

Le critère de sélection de l'espace couleur le plus discriminant est basé sur le nombre de pics significatifs des histogrammes  $h^{j,k}(x,\tau)$  puisque nous connaissons *a priori* le nombre  $N_C$  de classes qu'il faut extraire de la ROI de référence. Ainsi, le pouvoir discriminant  $R^{j,k}$  de la  $k^{\text{ème}}$  composante couleur du  $j^{\text{ème}}$  espace couleur est égal au nombre de pics significatifs  $N_P^{j,k}$  détectés de l'histogramme de la composante couleur, soit :  $R^{j,k} = N_P^{j,k}$ .

Deux cas différents peuvent apparaître pour déterminer l'espace couleur le plus discriminant J suivant que plusieurs espaces couleurs j possèdent au moins un histogramme lissé  $h^{j,k}(x,\tau)$  ayant  $N_C$  pics significatifs ou pas. Dans le cas ou plusieurs espaces couleurs j possèdent au moins un histogramme lissé  $h^{j,k}(x,\tau)$  ayant  $N_C$  pics significatifs, l'espace couleur J le plus discriminant est celui possédant le nombre le plus important d'histogramme lissé  $h^{j,k}(x,\tau)$  ayant  $N_C$  pics significatifs. Sinon, dans le cas ou aucun histogramme lissé  $h^{j,k}(x,\tau)$  ne possède  $N_C$  pics significatifs alors l'espace couleur J le plus discriminant est celui possédant l'histogramme lissé  $h^{j,k}(x,\tau)$  avec le nombre de pics significatifs le plus proche de  $N_C$ . Si plusieurs espaces couleur remplissent cette dernière condition, l'espace couleur J le plus discriminant correspond à celui dont la composante couleur avec le nombre de pics significatifs le plus proche de  $N_C$ , possède la plus grande somme des degrés de connexité  $CD(S_i^{j,k})$  (les degrés de connexité sont décrits dans le paragraphe 3.5.3.1).

Nous proposons d'illustrer la sélection de l'espace couleur *J* le plus discriminant à la première itération de la procédure pour la ROI de référence de la figure 5.10(b). Tous les pixels de la ROI de référence sont à analyser à cette itération. Le tableau 5.2 recense le nombre de pics significatifs des histogrammes lissés étudiés ainsi que la somme des degrés de connexité  $CD(S_i^{j,k})$ des sous-ensembles de pixels  $S_i^{j,k}$ . L'espace couleur le plus discriminant *J* sélectionné à cette itération est l'espace couleur (*Y*,*U*,*V*). En effet, bien que  $N_P^{6,3} = N_P^{8,2} = N_P^{11,2} = 3$ , la deuxième composante couleur de l'espace (*Y*,*U*,*V*) dont l'histogramme lissé possède le nombre de pics le plus proche de  $N_C$  est plus proche de la valeur  $N_C = 5$  avec  $N_P^{6,2} = 2$  que celles des deux autres espaces couleur dont les autres composantes ne possèdent qu'un seul pic significatif.

La figure 5.11 montre les histogrammes de la ROI de référence de la figure 5.10(b) pour les composantes couleur Y, U et V de l'espace (Y, U, V), ainsi que leurs histogrammes lissés pour la première itération de l'algorithme.

Une fois l'espace couleur *J* le plus discriminant sélectionné, celui-ci est empilé dans la *pile des espaces* qui recense les espaces couleur les plus discriminants utilisés pour la construction

j <sup>èm</sup>	<sup>e</sup> Espace couleur	$N_P^{j,1}$	$N_P^{j,2}$	$N_P^{j,3}$	$\sum_{i=1}^{N_P^{j,1}} CD(S_i^{j,1})$	$\sum_{i=1}^{N_P^{j,2}} CD(S_i^{j,2})$	$\sum_{i=1}^{N_P^{j,3}} CD(S_i^{j,3})$
1	(R,V,B)	1	1	1	0.986	0.972	0.976
2	(r,v,b)	2	2	2	1.887	1.744	1.832
3	(X,Y,Z)	1	1	1	0.972	0.971	0.977
4	(x, y, z)	2	1	2	1.886	0.986	1.833
5	(Y, I, Q)	1	2	2	0.970	1.889	1.833
6	(Y, U, V)	1	2	3	0.970	1.831	2.573
7	( <i>I</i> 1, <i>I</i> 2, <i>I</i> 3)	1	2	2	0.971	1.890	1.752
8	(wb, rg, by)	1	3	1	0.970	2.573	0.980
9	$(Y,C_1,C_2)$	1	2	1	0.970	1.890	0.986
10	$(L^*,a^*,b^*)$	1	2	2	0.985	1.757	1.825
11	$(L^*,u^*,v^*)$	1	3	1	0.986	2.569	0.985

**Tableau 5.2 :** Calcul des pouvoirs discriminants à la première itération appliquée à la ROI de la fi-<br/>gure 5.10(b).



**Figure 5.11 :** Histogrammes des composantes couleur de l'espace (Y, U, V) (i = 6) à la première itération appliquée à la ROI de la figure 5.10(b).

d'une classe de pixels à chaque itération de l'algorithme de la figure 5.9. Cette pile des espaces est utilisée ultérieurement par la procédure en-ligne. Une fois l'espace couleur le plus déterminant sélectionné, celui-ci est alors utilisé par l'étape n°7 de l'algorithme de la figure 5.9 pour construire une classe de pixels.

#### 5.3.1.6 Construction d'une classe de pixels

La construction de la classe de pixels est analogue à celle présentée dans le paragraphe 4.6. Les seuils déterminés permettant de délimiter le domaine de définition de la classe de pixels construite sont empilés dans la *pile des seuils* qui recense les seuils utilisés pour la construction d'une classe de pixels à chaque itération. L'algorithme de la figure 5.9 se poursuit alors en étape  $n^{\circ}8$ .

#### 5.3.1.7 Extraction des pixels assignés à la classe construite

Les pixels dont les points couleur sont inclus dans le domaine de définition de la classe construite lors de l'étape précédente sont extraits de la ROI de référence afin de ne pas être pris en compte lors de l'itération suivante.

La figure 5.12 montre les résultats obtenus à la première itération de notre méthode appliquée à la ROI de référence de la figure 5.10(b). Les pixels étiquetés en rouge de l'image 5.12(a) correspondent aux pixels appartenant à la première classe extraite par la méthode proposée et l'image 5.12(b) montre les pixels restant à traiter pour les itérations suivantes. Les pixels assignés lors de la première itération de l'algorithme correspondent aux pixels de couleur orange foncé.



(a) Pixels appartenant à la classe construite lors de la première itération.



(b) Pixels considérés par la seconde itération.

Cette étape permet donc d'extraire de la ROI de référence les pixels assignés à la classe construite. Une fois l'itération terminée, il convient de déterminer si il faut construire une autre classe ou si l'algorithme doit s'arrêter.

#### 5.3.1.8 Critère d'arrêt

L'algorithme s'arrête lorsque  $N_C$  classes de pixels ont été extraites. La phase d'apprentissage est alors terminée.

La figure 5.13 montre les pixels pour la ROI de référence de la figure 5.10(b) assignés aux classes construites lors des itérations successives. Les pixels assignés aux classes ainsi

**Figure 5.12 :** Résultats de la première itération de l'algorithme de la figure 5.9 appliquée à la ROI de référence 5.10(b).

construites sont montrés sur les images 5.13(a), 5.13(b), 5.13(c), 5.13(d) et 5.13(e) pour respectivement les itérations de rang 1, 2, 3, 4 et 5. Cette figure montre que notre méthode a réussi à discriminer les cinq classes de pixels présentes dans la ROI de référence.



(a) Pixels assignés à la classe construite lors de la première itération.



(d) Pixels assignés à la classe construite lors de la quatrième itération.



(b) Pixels assignés à la classe construite lors de la deuxième itération.



(e) Pixels assignés à la classe construite lors de la cinquième itération.

Figure 5.13 : Résultats de la ROI "Fanta".

Après avoir détaillé les différentes étapes de l'algorithme de la figure 5.9, nous allons maintenant présenter la procédure en-ligne qui exploite les résultats fournis par la procédure horsligne pour déterminer si le verre à inspecter est conforme.

### 5.3.2 Procédure en-ligne

L'algorithme de la procédure en-ligne est présenté sur la figure 5.14. Celui-ci est divisé en 8 étapes différentes qui seront détaillées dans les paragraphes suivants.



(c) Pixels assignés à la classe construite lors de la troisième itération.



(f) Image segmentée.



Figure 5.14 : Algorithme de la procédure en-ligne.

#### 5.3.2.1 Acquisition d'un verre à inspecter

La première étape consiste à faire l'acquisition de l'image représentant la surface du verre inspecté. La seconde étape de l'algorithme de la figure 5.14 consiste à rechercher les ROI déterminées lors de la phase d'apprentissage dans l'image à analyser. La recherche de ces ROI est expliquée dans le paragraphe suivant.

#### 5.3.2.2 Recherche des ROI

Cette recherche des ROI est effectuée par un outil de corrélation normalisée disponible dans la librairie Evision fournie par la société Euresys [Eur05b]. Cet outil compare la ROI de référence et une zone de l'image considérée. Il retourne alors un score compris entre 0 et 1, où le score 1 signifie une ressemblance parfaite entre la ROI de référence et la zone considérée tandis qu'un score de 0 signifie qu'il n'y a aucune ressemblance entre ces deux zones. Un seuil sur le score retourné par l'outil de corrélation normalisé permet alors de déterminer si la zone examinée correspond à une zone du verre que l'on souhaite étudier par maximum de ressemblance.

A l'issue de cette recherche, deux cas peuvent se produire suivant que la ROI maître est détectée ou pas dans l'image considérée. Si aucune ROI maître n'est détectée en accord avec le seuil fixé par l'utilisateur, alors le verre est considéré comme rebut puisqu'aucune zone du verre ne correspond à la ROI de référence. L'algorithme de la procédure en-ligne de la figure 5.14 s'arrête. Dans le cas où la ROI maître est retrouvée dans l'image du verre inspecté, les décalages  $\Delta x$  et  $\Delta y$  déterminés lors de la phase d'apprentissage sont appliqués pour localiser la ROI esclave dans l'image du verre à inspecter. Suivant que le nombre  $N_C$  de classes est égal ou supérieur à 1, l'algorithme poursuit respectivement en étape n°3 ou n°5. Le paragraphe 5.3.2.3 présente le cas où  $N_C = 1$  et le paragraphe 5.3.2.4 présente le cas où  $N_C > 1$ .

#### 5.3.2.3 Cas mono-modal

#### Calcul des vecteurs gradients des composantes R, V et B

Comme une seule classe de pixels doit-être présente dans la ROI examinée, les couleurs des pixels doivent être homogènes. Pour vérifier cela, nous appliquons un gradient couleur afin de détecter d'éventuelles transitions locales de couleurs. Le module du gradient appliqué est celui de Di-Zenzo [Zen86].

#### Fusion des 3 vecteurs gradients

Le module du gradient est seuillé avec une valeur faible afin que l'algorithme soit sensible à la moindre hétérogénéité locale de couleurs indiquant la présence d'un défaut d'aspect.

#### 5.3.2.4 Cas multi-modal

Si le nombre  $N_C$  de classes au sein de la ROI de référence est supérieur à 1 alors l'algorithme de la figure 5.14 segmente par multi-seuillage la ROI esclave de l'image du verre à inspecter avec les informations provenant de la pile des espaces et de la pile des seuils déterminée lors de la phase d'apprentissage. Comme la ROI esclave de l'image du verre de référence et celle de l'image du verre à inspecter sont similaires, nous supposons que les espaces couleur les plus discriminants et les seuils déterminés lors de la phase d'apprentissage sont toujours pertinents pour l'analyse de la ROI esclave du verre à inspecter. Les paragraphes suivants détaillent les différentes étapes de l'algorithme 5.14 pour l'étude d'une ROI esclave contenant plusieurs classes de pixels ( $N_C > 1$ ).

#### Conversion de la ROI dans l'espace couleur le plus discriminant

Pour chaque classe de pixels à construire, l'espace couleur le plus discriminant a été empilé dans la pile des espaces lors de la phase d'apprentissage. Il suffit alors de dépiler l'élément en tête de pile pour connaître l'espace couleur le plus discriminant à chaque itération. Les couleurs des pixels de la ROI esclaves sont codés dans cet espace.

#### Construction et extraction d'une classe de pixels

Une fois la ROI esclave de l'image du verre à inspecter convertie dans l'espace couleur le plus discriminant, la pile des seuils déterminés lors de la phase d'apprentissage fournit les seuils à appliquer pour construire la classe de pixels à chaque itération. Ces seuils sont ainsi dépilés de la pile des seuils et permettent de délimiter le domaine de définition classe de pixels pour chaque itération.

Les pixels dont les points couleur sont inclus dans le domaine de définition de la classe sont extraits de la ROI esclave à inspecter afin de ne pas être pris en compte lors de l'itération suivante. Les pixels assignés à cette classe sont comparés à l'étape n°7 de l'algorithme de la figure 5.14 avec les pixels appartenant à la classe de la même itération pour la ROI de référence.

#### Comparaison des pixels assignés aux classes

Lors de cette étape, l'algorithme compare les pixels de la ROI de référence à ceux de la ROI esclave qui ont été assignés à la même classe.

Une soustraction d'images permet de vérifier si ces pixels sont spatialement répartis de manière identique dans les deux ROI. Cependant, à cause du bruit d'acquisition, des déformations géométriques des verres, du décalage entre la ROI maître et la ROI esclave, ces pixels ne sont pas exactement en vis à vis. Pour cette raison, on recale les ROI esclaves de l'image de référence et de l'image à inspecter de manière à ce que le nombre de pixels en regard entre les deux ROI soit le plus important possible. Pour éviter que les effets de bord ne soit considérés comme défauts, on applique une ouverture avec un élément structurant de forme carrée et de dimension  $5 \times 5$  pixels. Nous appelons l'image résultante *image de décision* puisqu'elle sera utilisée pour déterminer si la ROI esclave analysée possède ou pas des pixels représentant un défaut.

La figure 5.15 illustre ces différentes opérations avec les pixels extraits de la ROI de référence de la figure 5.10(b) lors de la première itération (c.f. l'image 5.15(a)). Les pixels d'une ROI esclave provenant de l'image d'un verre à inspecter assignés à la même classe sont montrés sur l'image de la figure 5.15(b). L'image 5.15(c) est l'image résultante de la soustraction des images 5.15(a) et 5.15(b). Dans l'image 5.15(c), les coordonnés des pixels étiquetés en blanc sont celles des pixels de la ROI de référence et de la ROI esclave qui n'ont pas été assignés à la même classe lors de la première itération. Ces pixels sont susceptibles de représenter un défaut dans la ROI inspectée. Dans le cas présent, ces pixels étiquetés en blanc sont uniquement dus aux effets de bord puisque les contenus des deux ROI sont visiblement identiques. Enfin, l'image 5.15(d) montre l'image 5.15(c) après le filtrage morphologique. On peut remarquer sur cette image de décision que les pixels étiquetés en blanc provenant du problème d'effet de bord ont été éliminés.

Une fois ce traitement effectué, il convient de décider si la classe analysée est conforme à la classe provenant du verre de référence. Cette décision est effectuée lors de l'étape n°8 de l'algorithme de la figure 5.14.



(a) Pixels dont les points assignés (b) Pixels assignés à la classe à la classe construite lors de la construite lors de la première itépremière itération pour la ROI de ration pour la ROI esclave du référence de la figure 5.10(b). verre de référence.



(c) Résultat de la soustraction (d) Réponse de l'image de la entre l'image de la figure 5.15(a) figure 5.15(c) après l'opération et l'image de la figure 5.15(b). d'ouverture.

Figure 5.15 : Comparaison des ROI.

#### 5.3.2.5 Décision

L'objectif de cette étape est de décider si la ROI esclave est conforme, à savoir si elle ne contient pas de défaut d'aspect en la comparant à la ROI de référence. Pour cela, cette étape retourne au système de vision le message "conforme" si aucun défaut n'est détecté sur la ROI analysée et le message "rebut" dans le cas échéant. Selon le nombre de classes à construire, les règles de décision sont différentes.

Pour établir la règle de décision dans le cas où la ROI possède une seule classe de pixels  $(N_C = 1)$ , nous supposons que le système visuel humain est très sensible au moindre défaut. Pour cette raison, on considère que la ROI esclave du verre à inspecter possède un défaut si au moins un pixel contour y est détecté.

Dans le cas où la ROI de référence possède plusieurs classes de pixels ( $N_C > 1$ ), une analyse en composantes connexes est d'abord effectuée sur l'image de décision afin d'en déterminer les régions. La zone du verre inspectée est alors considérée comme défectueuse si au moins une région est constituée d'au moins  $S_P$  pixels et le message "rebut" est retourné au système de vision. Si aucune région n'est formée par au moins  $S_P$  pixels alors le message "conforme" est retourné. La règle de décision dépend du seuil  $S_P$  sur la taille des régions. Nous pensons qu'il est important que ce paramètre soit ajustable parce que l'œil humain est plus où moins sensible à certaines couleurs et à la complexité de la zone du verre étudiée. Ainsi, nous supposons dans ce travail que seul un opérateur averti peut fixer ce paramètre au cours de la phase d'apprentissage.

Après avoir décrit la procédure hors-ligne et la procédure en-ligne, le paragraphe suivant montre quelques résultats obtenus en terme de détection de défauts.

# 5.4 Résultats expérimentaux

La méthode proposée a été appliqué à l'inspection de 29 verres possédant un nombre de classes supérieur à un ( $N_C > 1$ ) et à l'inspection de 52 verres mono-colorés ( $N_C = 1$ ). Pour l'expérimentation, le seuil  $S_P$  a été fixé à 50 pixels.

La figure 5.16 montre quelques exemples de résultats obtenus quand les couleurs des pixels de la ROI inspectée sont homogènes ( $N_C = 1$ ). Pour chaque ligne, l'image de gauche est la ROI esclave analysée et l'image de droite montre les pixels contours détectés par le seuillage du module du gradient Di Zenzo. Dans cette figure, seule la première ROI analysée ne présente aucun défaut. Cette figure montre que les résultats obtenus sont très satisfaisants car tous les défauts sont détectés.



**Figure 5.16 :** Contrôle d'aspect d'une ROI esclave avec une seule couleur homogène ( $N_c = 1$ ) par réponse à un gradient couleur.

La figure 5.17 montre un exemple de résultat obtenu par la méthode de multi-seuillage proposée lorsque cinq classes ( $N_C = 5$ ) sont à analyser.

L'image en haut à gauche de la figure 5.17 montre l'image de référence sans défaut et l'image juste en dessous montre ses pixels extraits à la première itération de l'algorithme de la figure 5.14. L'image en haut à droite de cette figure montre l'image de la zone d'un verre à inspecter ayant un défaut et l'image juste en dessous montre les pixels extraits pour la même itération. On peut voir que le défaut représenté par une région de 97 pixels étiquetés en blanc est bien détecté sur l'image résultante de la soustraction. Cette décision est conforme a celle de l'expert.



Figure 5.17 : Contrôle d'aspect d'une ROI esclave avec cinq classes de pixels ( $N_c = 5$ ) par la méthode de multi-seuillage proposée.

Afin d'évaluer les résultats retournés par la méthode proposée, nous les avons confrontés à l'avis d'un expert. Pour estimer ces résultats, nous distinguons trois cas possibles suivant que l'avis de l'expert et le résultat retourné par la méthode sont concordants ou non :

- Bonne détection : l'expert qui a jugé la zone inspectée du verre et la méthode de contrôle d'aspect fournissent la même décision.
- Sur-détection : la zone inspectée du verre est "conforme" pour l'expert alors que notre méthode retourne le résultat "rebut".
- Sous-détection : la zone inspectée du verre est considéré comme "rebut" pour l'expert alors que notre méthode retourne le résultat "conforme".

Les résultats obtenus sont répertoriés sous forme de pourcentage de bonnes-détections, de sur-détection ou de sous-détection dans le tableau 5.3. Ce tableau montre que les performances

	Bonne détection	Sur-détection	Sous-détection
ROI où $N_C = 1$	93%	0%	7%
ROI où $N_C = 5$	88%	6%	6%

Tableau 5.3 : Résultats obtenus par notre méthode appliquée aux images.

atteintes par notre méthode dans les deux cas sont très élevées. En effet, dans une grande majorité des cas, notre système répond de manière analogue à l'expert. De plus, nous pensons que le pourcentage de bonnes détections peut encore être amélioré en analysant un nombre plus important de ROI. En effet, la zone du verre inspectée présente parfois le défaut présent sur la ROI esclave analysée de manière plus prononcée dans une zone voisine.

# **Conclusion générale**

Ce document décrit en détail notre contribution à la segmentation d'images couleur et présente les différentes étapes qui ont été nécessaires au développement de notre approche. La méthode que nous proposons s'intègre dans un système de contrôle qualité de l'aspect de motifs sérigraphiés sur des verres. Ce système est composé d'une caméra linéaire couleur qui assure l'acquisition d'images-ligne couleur successives. Les images-ligne ainsi acquises sont juxtaposées pour former l'image matricielle du développé de la surface du verre inspecté. Ces images matricielles sont alors segmentées et les pixels avec la même étiquette correspondent à une même classe de pixels. Les pixels connexes avec les mêmes étiquettes forment des régions qui seront exploitées par un algorithme de décision pour classer les images en deux catégories : images de motifs sérigraphiés de bonne qualité et images de motifs présentant des défauts d'aspect. L'idée fondamentale de notre approche est d'analyser la couleur des pixels afin de différencier ceux qui représentent le motif de ceux qui représentent un défaut.

Durant nos travaux de thèse, nous avons mis au point un dispositif d'inspection dont l'élément principal est une caméra linéaire couleur. Par conséquent, nous avons insisté dans le premier chapitre sur les techniques et géométries d'éclairage mise en oeuvre pour l'acquisition d'images-lignes au moyen d'une caméra linéaire couleur. Ces techniques consistent à obtenir une lumière incidente la plus homogène possible sur la ligne de visée de la caméra sans zones d'ombre ni reflets. Par ailleurs, nous avons choisi une source de lumière blanche qui permet d'exciter les capteurs photo-sensibles R, V et B de la caméra dans les mêmes proportions et restituer ainsi toutes les couleurs des motifs du verre à inspecter. Nous nous sommes ensuite particulièrement intéressés au fonctionnement des caméras linéaires couleur utilisant trois rangées de capteurs R, V et B montées en parallèle car cette technologie est la plus répandue pour des applications industrielles. L'emploi de ce type de caméra nécessite le respect de règles technologiques, physiques et optiques sous peine de faire apparaître des fausses couleurs dans les images-ligne acquises. Ce chapitre nous a permis ainsi d'appréhender l'acquisition d'images couleur représentant les motifs d'un décor imprimé par sérigraphie sur des verres cylindriques.

Les mémoires images matricielles couleur accessibles par l'utilisateur sont généralement codées dans un espace (R, V, B) puisqu'elles sont issues d'une caméra linéaire couleur dont les éléments photo-sensibles sont recouverts de filtres rouge, vert et bleu. Cependant, il existe une multitude d'autres espaces pour représenter la couleur et il est admis que le choix de l'espace

couleur est un point crucial pour le traitement des images couleur comme la segmentation. C'est pourquoi, nous avons été amenés lors du second chapitre à décrire les espaces couleur les plus fréquemment utilisés en segmentation d'images couleur. Ce chapitre aboutit à la conclusion que cette multitude d'espaces couleur rend difficile la détermination d'un espace couleur le mieux adapté à un problème spécifique de segmentation.

Aussi, nous tentons d'apporter un élément de réponse à ce problème en nous plaçant dans un contexte de segmentation par classification non supervisée de pixels. Pour ce faire, nous nous sommes dirigé vers une segmentation d'images par multi-seuillage des composantes couleur d'un espace qui a été décrite au troisième chapitre. Le produit cartésien des seuils détectés selon chacune des composantes permet de définir des parallélépipèdes rectangles. L'ensemble de ces parallélépipèdes sont examinés pour déterminer ceux qui constituent les domaines de définition des classes de pixels dans l'espace couleur considéré. La détection des seuils s'effectue par l'analyse des histogrammes mono-dimensionnels des trois composantes couleur associées à l'espace couleur. Le lissage de chacun des histogrammes mono-dimensionnels afin de mettre en évidence ses pics significatifs ainsi que la détection des seuils délimitant ces pics sont les deux points clés de la segmentation d'images couleur par multi-seuillage.

Nous avons par conséquent proposé une approche originale, nommée "spatial scale-space filtering", qui conjointement filtre l'histogramme mono-dimensionnel d'une composante couleur et détecte les seuils délimitant les pics significatifs en prenant en compte la répartition spatiale des pixels dont les niveaux sont inclus entre ces seuils. L'application de cette approche a mis en évidence l'amélioration des résultats par rapport à ceux obtenus par l'approche classique du "scale space filtering".

Le choix de l'espace couleur dans lequel sont codées les couleurs des pixels traités par un algorithme de segmentation par multi-seuillage est un problème actuellement ouvert. Dans le quatrième chapitre, nous apportons notre contribution en proposant une méthode itérative de segmentation. Sa principale originalité est la sélection non supervisée de l'espace couleur le plus discriminant pour construire chacune des classes de pixels en présence dans l'image.

A chaque itération, l'algorithme compare les pouvoirs discriminants des espaces couleur pour sélectionner l'espace le plus discriminant. L'algorithme détermine alors un domaine de définition d'une classe ayant la forme d'un parallélépipède rectangle dans cet espace. Les pixels dont les points couleur sont inclus dans ce domaine sont assignés à cette classe et extraits de l'image à segmenter pour ne plus être considérés lors des itérations suivantes.

Un soin particulier a été apporté à la définition du pouvoir discriminant d'un espace couleur basé sur la somme des degrés de connexité des sous-ensembles de pixels associés aux pics significatifs des histogrammes mono-dimensionnels. Ce pouvoir discriminant se révèle être efficace pour estimer dans quelle mesure la classe de pixels construite à chaque itération correspond à une région significative dans l'image. Cette mesure prend simultanément en considération les distributions des points dans l'espace couleur considéré et l'arrangement spatial des pixels caractérisés par ces couleurs dans l'image. La version actuelle de notre algorithme fournit des résultats prometteurs sur des images de référence.

Cette approche a été adaptée et intégrée dans un système industriel de contrôle de contrôlequalité de l'aspect des motifs de verres sérigraphiés. Nous avons montré au cinquième et dernier chapitre comment notre approche s'applique à la segmentation d'images-ligne couleur représentant la surface de verres sérigraphiés afin de détecter l'apparition d'éventuels défauts d'aspect. Ce chapitre a montré que notre méthode permet d'obtenir des résultats de détection de défauts très satisfaisants. Ceci montre que la détermination de l'espace couleur pour construire chacune des classes de pixels permet une discrimination des pixels représentant les motifs sérigraphiés de ceux représentant des défauts d'aspect.

Les résultats prouvent également la faisabilité méthodologique du contrôle de l'aspect de motifs couleur sérigraphiés sur des verres cylindriques. Le projet industriel doit maintenant se poursuivre en adaptant le poste prototype d'inspection réalisé au sein de l'EIPC afin de l'installer sur une ligne de production. Cette phase nécessitera d'abord de quantifier les temps de calcul des différentes procédures logicielles mises en oeuvre, puis d'optimiser leur implémentation afin de respecter les contraintes en terme de temps de réponse du système spécifiées par l'industriel.

Ayant consacré une période importante de ma thèse de doctorat à la réalisation du poste prototype d'inspection, nous n'avons pas exploré toutes les améliorations possibles de notre algorithme de segmentation par construction itérative des classes de pixels. En effet, les étapes de lissage des histogrammes mono-dimensionnels et de détection des seuils délimitant les pics significatifs méritent d'être affinées. Par ailleurs, nous sommes conscients que la mesure du pouvoir discriminant d'un espace couleur basée sur les degrés de connexité gagnerait à prendre en compte le caractère tri-dimensionnel de l'information couleur [MVP06].

De même le critère actuellement utilisé sélectionne parfois des classes correspondant à du bruit et des contours. Ainsi, la détermination d'une classe de pixels à extraire à chaque itération peut sans doute être améliorée en prenant conjointement en compte les degrés de connexité et la population des pixels.

Enfin, les méthodes que nous avons développées se basent sur l'hypothèse suivante : les couleurs des régions à trouver dans les images sont homogènes. Or les régions présentes dans de nombreuses images ne respectent pas cette hypothèse, et notamment celles présentant des textures couleur. Ce constat est également une réalité industrielle puisqu'une quantité relativement importante de verres sérigraphiés produits présentent des motifs texturés susceptibles d'avoir des défauts d'aspect. Les procédures logicielles que nous avons mises au point atteignent leur limites pour la détection de défauts apparaissant dans des motifs couleur texturés.

Il faudra alors poursuivre le travail accompli par l'analyse de textures couleur présentes

dans les images numériques. Peu de travaux sont parus sur les liens entre la caractérisation de la texture couleur et les espaces de représentation de la couleur qui semblent constituer une voie de recherche prometteuse, apportant des perspectives d'amélioration de la segmentation d'images couleur.

# Bibliographie

- [AFN75] AFNOR. Dictionnaire de colorimétrie théorique et technique. Technical report, AFNOR, Décembre 1975.
- [AP06] S. Arlt et K. Pahnke. emsys vision ip cores, http://www.emsys.de/news/vision\_flyer\_en.pdf, 2006.
- [AR96] I. S. Ahmad et J. F. Reid. Evaluation of color representations for maize images. *Journal of Agricultural Engineering Research*, 63(3):185–195, 1996.
- [Ass04] Automated Imaging Association. Specifications of the camera link interface standard for digital cameras and frame grabbers. Technical Report Version 1.1, Automated Imaging Association, 2004.
- [ATM06] ATMEL. Camera Link Color Linescan Camera AVIIVA SC2CL. ATMEL, Juin 2006.
- [BAS03] BASLER. BASLER L301-kc User's Manual. Basler Vision Technologies, 2003.
- [BB97] J.-P. Braquelaire et L. Brun. Comparison and optimization of methods of color image quantization. *IEEE transactions on image processing*, 6(7) :1048–1052, Juillet 1997.
- [BCS98] M. Borsotti, P. Campadelli, et R. Schettini. Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19(8) :741–747, 1998.
- [Bez81] J.C. Bezdeck. Pattern recognition with fuzzy objective function algorithms. *Plenum Press*, 1981.
- [BMB<sup>+</sup>92] C. C. Brunner, A. G. Maristany, D. A. Butler, D. Vanleuween, et J. W. Funck. An evaluation of color spaces for detecting defects in Douglas-fir veneer. *Industrial Metrology*, 2(3 and 4) :169–184, 1992.
- [Bor04] H. Le Borgne. *Analyse de scènes naturelles par composantes indépendantes*. Thèse de doctorat, Institut National Polytechnique de Grenoble, Janvier 2004.
- [Bru02] L. Brun. *Traitement d'images couleur et Pyramides combinatoires*. Habilitation à diriger des recherches, Université de Reims, Décembre 2002.
- [BVMP04a] L. Busin, N. Vandenbroucke, L. Macaire, et J.-G. Postaire. Color space selection for unsupervised color image segmentation by analysis of connectedness properties. In *First Complex Systems Intelligence and Modern Technological Applications*, Septembre 2004.
- [BVMP04b] L. Busin, N. Vandenbroucke, L. Macaire, et J.-G. Postaire. Color space selection for unsupervised color image segmentation by histogram multithresholding. In *IEEE International Conference on Image Processing*, pages 203–206, Octobre 2004.
- [BVMP05] L. Busin, N. Vandenbroucke, L. Macaire, et J.-G. Postaire. Color space selection for unsupervised color image segmentation by analysis of connectedness properties. *International Journal of Robotics & Automation*, 20(2) :70–77, 2005.

[BVMP06]	L. Busin, N. Vandenbroucke, L. Macaire, et JG. Postaire. Scale space filter based on homogeneity degree for color image segmentation. In <i>Colour in Graphics</i> <i>Imaging and Vision</i> , pages 12–17. IS&T, Juin 2006.
[BW02]	B. G. Batchelor et P. F. Whelan. <i>Intelligent Vision Systems for Industry</i> . Springer Verlag, 2002.
[Car95]	T. Carron. Segmentation d'images couleur dans la base teinte-luminance- saturation : approche numérique et symbolique. Thèse de doctorat, Université de Savoie, Décembre 1995.
[Car00]	J. Caron. Les systèmes de vision linéaire pour le contrôle qualité dans les procé- dés de production continue. Thèse de doctorat, Université de Lille 1, Décembre 2000.
[Cel90]	M. Celenk. A color clustering technique for image segmentation. <i>Computer Vision, Graphics and Image Processing</i> , 52(2):145–170, 1990.
[CH04]	Y. Chen et P. Hao. Optimal transform in perceptually uniform color space and its application in image retrieval. In <i>International Conference on Signal Processing</i> , pages 1107–1110, 2004.
[CIE86]	Commission Internationale de l'Eclairage. Colorimetry. Technical Report 15.2, Bureau central de la CIE, Vienna, 1986.
[CJSW01]	HD. Cheng, XH. Jiang, Y. Sun, et J. Wang. Color image segmentation : advances and prospects. <i>Pattern Recognition</i> , 34 :2259–2281, 2001.
[CJW02]	H.D. Cheng, X. H. Jiang, et J. Wang. Color image segmentation based on homogram thresholding and region merging. <i>Pattern Recognition</i> , 35(2):373–393, 2002.
[CKHD <sup>+</sup> 97]	C. Coulot, S. Kohler-Hemmerlin, C. Dumont, D. Aluze, et B. Lamalle. Lighting study for an optimal defects detection by artificial vision. <i>Machine Vision Applications in Industrial Inspection V</i> , 3029 :69–77, Février 1997.
[CL03]	H. D. Cheng et J. Li. Fuzzy homogeneity and scale-space approach to color image segmentation. <i>Pattern Recognition</i> , 36 :1545–1562, 2003.
[CM02]	D. Comaniciu et Peter Meer. Mean shift : A robust approach toward feature space analysis. <i>IEEE transactions on Pattern Analysis and Machine Intelligence</i> , 24(5), Mai 2002.
[Cou91]	V. Coutance. <i>La couleur en vision par ordinateur</i> . Thèse de doctorat, Université Paul Sabatier de Toulouse, Janvier 1991.
[CP95]	JP. Coquerez et S. Philipp. <i>Analyse d'images : filtrage et segmentation</i> . Masson, 1995.
[CS00]	H.D. Cheng et Y. Sun. A hierarchical approach to color image segmentation using homogeneity. <i>IEEE Trans. on Image Processing</i> , 9(12) :2071–2082, 2000.
[DAL05]	DALSA. Trilium Camera User's Manual. DALSA, 2005.
[DMS99]	Y. Deng, B.S. Manjunath, et H. Shin. Color image segmentation. In <i>Proc. of the IEEE Computer Vision and Pattern Recognition - CVPR'99</i> , pages 446–451, Fort Collins, USA, 1999.
[DS05]	CJ. Du et DW. Sun. Comparison of three methods for classification of pizza topping using different colour space transformations. <i>Journal of Food Engineering</i> , 68 :277–287, 2005.

[ES05]	E.Simonetto et A. Saadane. Colour image watermarking using a visual sub-band decomposition. <i>International Journal of Robotics &amp; Automation</i> , 20 :101–108, 2005.
[EUR03]	EUROSEP. <i>Générateur de lumière froide SPOT.IL.5050</i> . Eurosep instruments, 2003.
[Eur05a]	Euresys. About ADR technology. Technical report, Euresys, 2005.
[Eur05b]	Euresys. eVision user's guide. Technical report, Euresys, 2005.
[Fau76]	O. D. Faugeras. <i>Digital color image processing and pshychophysics within the framework of a human visual model</i> . Thèse de doctorat, University of Utah, Juin 1976.
[Fau79]	O. D. Faugeras. Digital color image processing within the framework of a human visual model. <i>IEEE Transactions on Acoustics, Speech, and Signal processing</i> , 27(4):380–393, 1979.
[Fir97]	C. Firmin. Optimisation des réseaux de neurones à fonctions radiales de base par critères informationnels : application à la détection de défauts en production de bouteilles. Thèse de doctorat, Université Lille 1, Janvier 1997.
[Fon01]	M. Fontaine. Segmentation non supervisée d'images couleur par analyse de la connexité des pixels. Thèse de doctorat, Université de Lille 1, Décembre 2001.
[FYEA01]	J. Fan, D. K. Y. Yau, A. K. Elmagarmid, et W.G. Aref. Automatic image seg- mentation by integrating color-edge extraction and seeded region growing. <i>IEEE</i> <i>Trans. on Image Processing</i> , 10(10) :1454–1466, 2001.
[Gar79]	C. Garbay. <i>Modélisation de la couleur dans le cadre de l'analyse d'images et de son application à la cytologie automatique</i> . Thèse de doctorat, Institut National Polytechnique de Grenoble, Décembre 1979.
[Gau06]	J. Gaudin. Colorimétrie appliquée à la vidéo. Dunod, 2006.
[GD03]	K. L. Gunther et K. R. Dobkins. Independence of mechanisms tuned along car- dinal and non-cardinal axes of color space : evidence from factor analysis. <i>Vision</i> <i>Research</i> , 43 :683–696, 2003.
[Gev00]	T. Gevers. Color in image databases. Technical report, University of Asmterdam, 2000.
[GH92]	J. Gauch et CW. Hsia. A comparison of three color image segmentation algorithm in four color spaces. <i>SPIE Visual Communications and Image Processing</i> , 1818 :1168–1181, 1992.
[Gou05]	M. Gouiffès. Apports de la couleur et des modèles de réflexion pour l'extraction et le suivi de primitives. Thèse de doctorat, Université de Poitiers, Décembre 2005.
[Gro04]	C. Groeppelin. Où va la vision? J'automatise, 33:82-85, Mars-Avril 2004.
[GSD01]	T. Géraud, PY. Strub, et J. Dardon. Color image segmentation based on automatic morphological clustering. In <i>IEEE International Conference on Image Processing</i> , pages 70–73, Thessalonique, Grêce, 2001.
[Han02]	A. Hanbury. <i>Morphologie mathématique sur le cercle unité avec applications aux teintes et aux textures orientéees</i> . Thèse de doctorat, Ecole Nationale des Mines de Paris, Mars 2002.
[Hel66]	H. Helmoltz. <i>Handbuch der physiologishen optik</i> . publisher, Hambourg et Leib- zig, 1866.

[Hub05]	P. M. Hubel. Foveon technology and the changing landscape of digital camera technology. In <i>Keynote paper at the thirteenth IS&amp;T Color Imaging Conference</i> , pages 314–317, 2005.
[IMA05a]	IMASYS. Chapitre 1. Eclairages. Technical report, Imasys, http://www.imasys.fr/technologies/1_eclairages_v1.pdf, Janvier 2005.
[IMA05b]	IMASYS. Chapitre 2. Optiques. Technical report, Imasys, http://www.imasys.fr/technologies/2_optiques_v101.pdf, Mai 2005.
[Ken76]	J. Kender. Saturation, hue and normalized color : Calculation, digitization effects, and use. Technical report, Departement of Computer Science, Carnegie-Mellon University, 1976.
[Kod03]	Kodak. CCD image sensor noise sources. Technical report, Eastman Kodak Company, Janvier 2003.
[Kow90]	P. Kowaliski. Vision et mesure de la couleur. Masson, seconde edition, 1990.
[KSH00]	F. Kurugollu, B. Sankur, et A. E. Harmanci. Color image segmentation using 2-dimensional histogram multithresholding and fusion. In <i>First International Conference on Color in Graphics and Image Processing - CGIP'2000</i> , pages 152–157, Saint-Etienne, France, 2000.
[LC99]	P. Lambert et T. Carron. Symbolic fusion of luminance-hue-chroma features for region segmentation. <i>Pattern Recognition</i> , 32(11):1857–1872, 1999.
[LCK94]	J. H. Lee, B. H. Chang, et S. D. Kim. Comparison of colour transformations for image segmentation. <i>Electronics Letters</i> , 30(20):1660–1661, 1994.
[Leb02]	V. Lebrun. Caractérisation des pierres ornementale par analyse d'images. Mas- ter's thesis, Université de Liège - Faculté des Sciences Appliquées, Mars 2002.
[LH93]	H. Levkowitz et G. T. Herman. GLHS : A generalized lightness, hue, and satura- tion color model. <i>CVGIP : Graphical Models and Images Processing</i> , 55(4) :271– 285, Juillet 1993.
[Lin90]	T. Lindeberg. Scale-space for discrete signals. <i>IEEE Transactions of Pattern Analysis and Machine Intelligence</i> , 12(3):234–254, 1990.
[Lin94]	T. Lindeberg. <i>Scale-space theory in computer vision</i> . Kluwer Academic Publi- sher, 1994.
[Liu93]	J. Liu. Multiresolution colour image segmentation. Master's thesis, University of Saskatchewan, Janvier 1993.
[LL90]	Y. W. Lim et S. U. Lee. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. <i>Pattern Recognition</i> , 23(9):935–952, 1990.
[LLL06]	C. Li, C. Li, et M. R. Luo. Quality of LED based daylight simulators. In <i>Conference on Color in Graphics, Imaging, and Vision</i> , pages 247–250, Leeds, Juin 2006. IS&T.
[LM00]	P. Lambert et L. Macaire. Filtering and segmentation the specificity of color images. In <i>International Conference on Color in Graphics and Image Processing</i> , pages 57–71, Saint-Etienne, 2000.
[LM01]	V. Lebrun et L. Macaire. Aspect inspection of marble tiles by colour line scan cameras. In <i>International Conference on Quality Control by Artificial</i> , pages 403–408, Le Creusot, 2001.
[Loz98]	V. Lozano. <i>Contribution de l'analyse d'image couleur au traitement des images textile</i> . Thèse de doctorat, Univesité Jean Monnet, Janvier 1998.
-----------------------	--
[LR97]	E. Littmann et H. Ritter. Adaptive color segmentation. A comparison of neural and statistical methods. <i>IEEE Transactions on Neural Networks</i> , 8(1):175–185, 1997.
[LRS <sup>+</sup> 01]	AC. Legrand, E. Renier, P. Suzeau, F. Truchetet, P. Gorria, et F. Meriaudeau. Machine vision systems in metallurgy industry. <i>Journal of Electronic Imaging</i> <i>SPIE</i> , 10(1):274Ű282, Janvier 2001.
[LWY96]	HC. Lin, LL. Wang, et SN. Yang. Automatic determination of the spread parameter in gaussian smoothing. <i>Pattern Recognition Letters</i> , 17 :1247–1252, 1996.
[LY94]	J. Liu et YH. Yang. Ieee transaction multiresolution color image segmentation. <i>Pattern Analysis and Machine Intelligence</i> , 16(7):689–700, Juillet 1994.
[Mac42]	D. L. MacAdam. Visual sensitivities to color differences in daylight. <i>Journal of the Optical Society of America</i> , 32 :247–274, 1942.
[Mac43]	D. L. MacAdam. Specification of small chromaticity differences. <i>Journal of the Optical Society of America</i> , 33 :18–26, 1943.
[Mac67]	L. MacQueen. Some methods for classification and analysis of multivariate observations. <i>Proc. of the 5th Berkeley Symp.</i> , pages 281–297, 1967.
[Mac93]	L. Macaire. <i>Inspection d'états de surface par vision linéaire</i> . Thèse de doctorat, Université de Lille 1, Janvier 1993.
[Mac04]	L. Macaire. <i>Habilitation à diriger des recherches, Exploitation de la couleur pour la segmentation et l'analyse d'images</i> . Université de Lille,, 2004.
[Mah05]	M. Mahmoudi. Pour bien voir, il faut aussi le bon objectif. <i>Mesures</i> , 773 :51–53, Mars 2005.
[Mat00]	H. Mathieu. La chaîne de l'acquisition d'images. Technical report, INRIA, Décembre 2000.
[Méc05]	P. Méchin. Les composants FPGA s'invitent dans le traitement du signal. <i>Mesures</i> , 775 :54–56, Mai 2005.
[Muk02]	J. Mukherjee. MRF clustering for segmentation of color images. <i>Pattern Recog-</i> <i>nition Letters</i> , 23(8):65–78, 2002.
[MVP04]	L. Macaire, N. Vandenbroucke, et JG. Postaire. Segmentation d'images par classification spatio-colorimétrique des pixels. <i>Traitement du signal</i> , 21(5):423–437, 2004.
[MVP06]	L. Macaire, N. Vandenbroucke, et JG. Postaire. Color image segmentation by analysis of subset connectedness and color homogeneity properties. <i>Computer Vision and Image Understanding</i> , 102(1):105–116, Avril 2006.
[Nev77]	R. Nevatia. A color edge detector and its use in scene segmentation. <i>IEEE Transactions on Systems, Man, and Cybernetics</i> , 7(11):820–826, 1977.
[NJ95]	T. S. Newman et A. K. Jain. A survey of automated visual inspection. <i>Computer Vision and Image Understanding</i> , 61 :231–262, Mars 1995.
[OKS80]	Y. I. Ohta, T. Kanade, et T. Sakai. Color information for region segmentation. <i>Computer Graphics and Image Processing</i> , 13:222–241, 1980.

[OPR78]	R. Ohlander, K. Price, et DR. Reddy. Picture segmentation using a recursive region splitting method. <i>Computer Graphics and Image Processing</i> , 8:313–333, 1978.
[OSR04]	OSRAM. Appareillages - Lampes - Luminaires 2004/2005. OSRAM, 2004.
[PBC05]	S. L. Phung, A. Bouzerdoum, et D. Chai. Skin segmentation using color pixel classification : Analysis and comparison. <i>IEEE transactions on Pattern Analysis and Machine Intelligence</i> , 27(1), Janvier 2005.
[Per]	PerkinElmer. YD5000-Series Camera Instruction Manual. PerkinElmer.
[Pos98]	JG. Postaire. Contrôle de production par vision linéaire. <i>Techniques de l'Ingé- nieur, traité Mesures et Contrôle</i> , R 7 600 :1–15, 1998.
[PR02]	I. Philipp et T. Rath. Improving plant discrimination in image processing by use of different colour space transformations. <i>Computers and Electronics in Agriculture</i> , 32(35):1–15, 2002.
[PYL98]	S.H. Park, I.D. Yun, et S.U. Lee. Color image segmentation based on 3-D clustering : morphological approach. <i>Pattern Recognition</i> , 31(8) :1061–1076, 1998.
[Raf99]	G. Raffy. <i>Vision numérique couleur appliquée à l'évaluation de la qualité de carcasses de dindes</i> . Thèse de doctorat, Université Blaise Pascal - Clermont II, Mars 1999.
[Ros98]	P. L. Rosin. Determining local natural scales of curves. <i>Pattern Recognition Letters</i> , 19:63–75, 1998.
[Rou98]	G. Rougeron. <i>Problèmes liés à la couleur en synthèse d'images</i> . Thèse de docto- rat, Ecole Nationale Supérieure des Mines de Saint-Etienne, Janvier 1998.
[SA81]	A. Sarabi et J. K. Aggarwal. Segmentation of chromatic images. <i>Pattern Recog-</i> <i>nition</i> , 13(6):417–427, 1981.
[SB91]	M. J. Swain et D. H. Ballard. Color indexing. <i>International Journal of Computer Vision</i> , 7(1):11–32, 1991.
[SC05]	F. Y. Shih et S. Cheng. Automatic seeded region growing for color image segmentation. <i>Image and Vision Computing</i> , 23:877–886, 2005.
[Sch93]	R. Schettini. A segmentation algorithm for color images. <i>Pattern Recognition Letters</i> , 14:499–506, Juin 1993.
[Sèv96]	R. Sève. <i>Physique de la couleur : De l'apparence colorée à la technique colori-</i> <i>métrique</i> . Masson, 1996.
[Shi95]	TY. Shih. The reversibility of six geometric color spaces. <i>Photogrammetric Engineering and Remote Sensing</i> , 61(10) :1223–1232, 1995.
[SP97]	L. Shafarenko et M. Petrou. Automatic watershed segmentation of randomly textured color images. <i>IEEE Transactions on Image Processing</i> , 6(11) :1530–1543, 1997.
[SPK98]	L. Shafarenko, M. Petrou, et J. Kittler. Histogram-based segmentation in a perceptually uniform color space. <i>IEEE Transactions. on Image Processing</i> , 7(9):1354–1358, 1998.
[SS04]	M. Sezgin et B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. <i>Journal of Electronic Imaging</i> , 13(1):146–168, Janvier 2004.

[Taner]	PM. Tancrez. Un système de vision linéaire pour l'inspection et la caractéri- sation en temps réel de produits plats en défilement continu. Thèse de doctorat, Université Lille 1, 1989 Janvier.
[TAT04]	TATTILE. Tapix LCD 8000@14000 Reference Manual. TATTILE, 2004.
[TB97]	A. Trémeau et N. Borel. A region growing and merging algorithm to color seg- mentation. <i>Pattern Recognition</i> , 30(7) :1191–1203, 1997.
[TC00]	A. Trémeau et P. Colantoni. Region adjacency graph applied to color image seg- mentation. <i>IEEE Trans. on Image Processing</i> , 9(4) :735–744, 2000.
[TFMB04]	A. Trémeau, C. Fernandez-Maloigne, et P. Bonton. <i>Image numérique couleur - de l'acquisition au traitement</i> . Dunod, 2004.
[Tom92]	S. Tominaga. Color classification of natural color images. <i>Color Research and Application</i> , 17(4) :230–239, 1992.
[UA94]	T. Uchiyama et M.A. Arbib. Color image segmentation using competitive lear- ning. <i>IEEE Trans. on Pattern Analysis and Machine Intelligence</i> , 16(12):1197– 1206, 1994.
[Van00]	N. Vandenbroucke. Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football. Thèse de doctorat, Université des Sciences et Technologies de Lille 1, Décembre 2000.
[VM05]	N. Vandenbroucke et L. Macaire. Représentation de la couleur en analyse d'images. <i>Techniques de l'ingénieur</i> , S 7 602 :1–22, 2005.
[VMP98]	N. Vandenbroucke, L. Macaire, et JG. Postaire. Color pixels classification in an hybrid color space. In <i>IEEE International Conference on Image Processing</i> , pages 176–180, Chicago, 1998.
[VMP00]	N. Vandenbroucke, L. Macaire, et JG. Postaire. Color systems coding for color image processing. In <i>First International Conference on Color in Graphics and Image Processing</i> , volume 1, pages 180–185, Saint-Etienne, 2000.
[VMP03]	N. Vandenbroucke, L. Macaire, et JG. Postaire. Color image segmentation by pixel classification in adapted hybrid color space. application to soccer image segmentation. <i>Computer Vision and Image Understanding</i> , 90(2):190–216, 2003.
[VMVP04]	N. Vandenbroucke, L. Macaire, C. Vieren, et JG. Postaire. <i>Analyse de séquences d'images couleur de matchs de football pour le suivi des joueurs</i> , chapter 9.3 Application de l'analyse d'images couleur, pages 342–352. Dunod, 2004.
[Wai03]	T. Waillly. Dix critères pour bien choisir sa caméra intégrée. <i>Mesures</i> , 759 :59–62, 2003.
[Wit84]	A. P. Witkin. Scale-space filtering : a new approach to multi-scale description. <i>Image Understanding</i> , pages 79–95, 1984.
[You07]	T. Young. Lectures on natural philosophy, volume II. Johnson, London, 1807.
[Zan02]	ML. Zani. Bien éclairer sa pièce pour mieux la contrôler. <i>Mesures</i> , 742 :43–46, Février 2002.
[Zan05]	ML. Zani. Les cartes d'acquisition et de traitement d'images. <i>Mesures</i> , 771 :48–56, Janvier 2005.
[Zen86]	S. Di Zenzo. A note on the gradient of a multi-image. <i>Computer Graphics and Image Processing</i> , 33(1):116–125, 1986.

- [ZL98] D. Zugaj et V. Lattuati. A new approach of color images segmentation based on fusing region and edge segmentations outputs. *Pattern Recognition*, 31(2):105– 113, 1998.
- [Zuc76] S. W. Zucker. Region growing : childhood and adolescence. *Computer Graphics and Image Processing*, 5 :382–399, 1976.
- [ZW96] X. Zhang et B. A. Wandell. A spatial extension of CIELAB for digital color image reproduction. Society for Information Display Symposium Technical Digest, 27 :731–734, 1996.

# **Table des figures**

1.1	Chaîne d'acquisition d'un système de vision	15
1.2	Image d'un capteur linéaire couleur Kodak $KLI - 2113$ , disponible à l'adresse	
	http://www.kodak.com/US/en/dpq/site/SENSORS/name/KLI-2113_product	16
1.3	Images acquises par des caméras linéaire et matricielle couleur	17
1.4	Poste d'acquisition d'images-ligne	19
1.5	Réflexion spéculaire.	21
1.6	Réflexion parfaitement diffuse.	22
1.7	Réflexion mixte	22
1.8	Éclairage diffus et éclairage direct.	23
1.9	Éclairage bright-field	24
1.10	Éclairage co-axial	25
1.11	Éclairage par dôme	26
1.12	Répartition spectrale relative d'énergie de l'illuminant D65 normalisé par la	
	CIE	27
1.13	Lignes de visée d'une caméra tri-linéaire avec et sans système de prismes	32
1.14	Architecture fonctionnelle d'une caméra linéaire couleur	33
1.15	Mode de déclenchement "Free run edge-controlled" pour une caméra linéaire	
	couleur	34
1.16	Mode de déclenchement "Free run programmable" pour une caméra linéaire	
	couleur	34
1.17	Mode de déclenchement "ExSync edge-controlled" pour une caméra linéaire	
	couleur	34
1.18	Mode de déclenchement "ExSync level-controlled" pour une caméra linéaire	
	couleur	35
1.19	Mode de déclenchement "ExSync level-Programmable" pour une caméra li-	
	néaire couleur.	35
1.20	Lignes de visée des capteurs d'une caméra linéaire couleur	36
1.21	Phénomène de décalage spatial	37
1.22	images-ligne successives juxtaposées dans une image matricielle	38

1.23	Courbes de sensibilité spectrale du capteur linéaire couleur Kodak KLI-2113	
	(données constructeur).	39
1.24	Exemple de problème causé par la non perpendicularité entre la direction lon- gitudinale et transversale.	41
1.25	Acquisition d'une ligne noire d'une pièce animée d'un <b>mouvement rectiligne</b>	
	suivant la direction longitudinale par une caméra linéaire couleur.	42
1.26	Acquisition d'une ligne noire d'une pièce animée d'un <b>déplacement non rec-</b>	
	tiligne suivant la direction longitudinale par une caméra linéaire couleur.	43
1.27	Problème d'acquisition engendré avec un angle $\theta \neq 0$	44
1.28	niveaux des composantes couleur des pixels $p$ et $q$ de la figure 1.27	45
1.29	Problème d'acquisition engendré avec un angle $\psi \neq 0$	45
1.30	niveaux des composantes couleur des pixels $p$ et $q$ de la figure 1.29	46
1.31	Principe de fonctionnement du mode d'acquisition ADR	49
2.1	Cube des couleurs $(R, V, B)$	57
2.2	Diagramme de chromaticité $(x, y)$	58
2.3	Espace couleur $(L^*, a^*, b^*)$ . Cette image est disponible sur internet à l'adresse :	
	http://tecfa.unige.ch/lombardf/CPTIC/couleurs/couleur_ERAG/Pages/ch8p1.ht	т
		60
2.4	espace de coordonnées cylindriques.	65
2.5	Modèle hexagonal et hexagonal double.	67
2.6	Les familles d'espaces couleur.	72
3.1	Multi-seuillage par analyse d'histogrammes 1D des composantes couleur	86
3.2	Projection d'un pixel du plan image en un point couleur dans l'espace couleur	
	(R,V,B)	87
3.3	Histogramme 3D de l'image 3.3(a) dans l'espace $(R, V, B)$ .	87
3.4	Histogrammes 1D de l'image de synthèse de la figure 3.3(a)	88
3.5	Convolution de l'histogramme de la composante bleue $h^B(x)$ de l'image 3.3(a)	
	avec un noyau gaussien d'écart- type $\tau = 2$ .	90
3.6	Convolution de l'histogramme de la composante bleue $h^B(x)$ de l'image 3.3(a)	
	avec un noyau gaussien d'écart-type $\tau = 5$	90
3.7	Convolution de l'histogramme de la composante bleue $h^B(x)$ de l'image 3.3(a)	
	avec un noyau gaussien d'écart-type $\tau = 10. \dots \dots \dots \dots \dots \dots$	91
3.8	Caractéristiques des histogrammes.	92
3.9	Dérivée première $h_x^k(x, \tau)$ de l'histogramme $h^k(x, \tau)$ de la figure 3.8	93
3.10	Dérivée seconde $h_{xx}^k(x, \tau)$ de l'histogramme $h^k(x, \tau)$ de la figure 3.8	94

3.11	Image de synthèse avec les histogrammes des composantes $R$ , $V$ et $B$ et une	
	modélisation des histogrammes à partir des deux gaussiennes $g_1(x, 30)$ et	
	$g_2(x,30)$	95
3.12	Influence du choix des seuils pour une image en niveaux de gris	96
3.13	Erreur de classification suivant les valeurs de seuils utilisés	97
3.14	Influence du choix des seuils pour la segmentation de l'image couleur de la	
	figure 3.11(a)	98
3.15	Partitionnement du plan couleur $(R, V)$	99
3.16	Représentation de la classe construite dans l'espace $(R, V, B)$ de la région $\Re 1$	
	ainsi que des pixels de l'image dont les niveaux de couleur sont inclus dans le	
	parallélépipède associé à cette classe	101
3.17	Méthode de lissage global. Chaque étape est numérotée de 1 à 6	103
3.18	Méthode de lissage local. Chaque étape est numérotée de 1 à 6	105
3.19	Détection des valeurs sur l'axe des abscisses des seuils des histogrammes	
	$h^B(x,2)$ , $h^B(x,5)$ et $h^B(x,10)$ de l'image de synthèse. Les points en rouge	
	montrent la localisation de ces seuils	108
3.20	Représentation des minima locaux des histogrammes lissés de la composante	
	bleue de l'image de synthèse par des noyaux gaussiens d'écarts-type 2, 5 et	
	10 dans le scale-space et l'empreinte de l'histogramme $h^B(x)$	109
3.21	Empreintes des histogrammes 3.4(a) 3.4(b) et 3.4(c) de l'image 3.3(a)	110
3.22	Nœuds de l'histogramme $h^B(x)$	111
3.23	Noeuds des histogrammes des composantes couleur rouge 3.4(a), verte 3.4(b)	
	et bleue 3.4(c) de l'image de synthèse 3.3(a)	112
3.24	Nœuds actifs sélectionnés par le SSF de l'histogramme $h^B(x)$	113
3.25	Pixels de l'image de synthèse 3.3(a) dont les niveaux de bleu appartiennent	
	aux pics correspondant aux nœuds actifs sélectionnés par le critère clasique	
	du SSF	114
3.26	Nœuds actifs de l'histogramme $h^B(x)$ sélectionnés interactivement	115
3.27	Pixels de l'image de synthèse 3.3(a) dont les niveaux de bleu appartiennent	
	aux pics correspondant aux nœuds actifs sélectionnés de manière interactive	115
3.28	Degré de compacité des noeuds provenant de l'histogramme $h^B(x)$ de l'image	
	de synthèse 3.3(a)	119
3.29	Résultats de la segmentation de l'image de synthèse 3.3(a) avec deux critères	
	de sélection des nœuds actifs différents.	120
3.30	Image originale et image segmentée de l'image "house"	121
3.31	Image originale et image segmentée de l'image "jelly beans"	122
3.32	Image originale et image segmentée de l'image "hand"	122

4.1	Algorithme de segmentation d'images couleur proposé par construction itéra-	
	tive des classes de pixels.	126
4.2	Les espaces couleur utilisés dans l'approche proposé sont notés en rouge	128
4.3	Histogrammes $h^{1,1}(x)$ , $h^{1,2}(x)$ et $h^{1,3}(x)$ des composantes rouge, verte et bleue	
	de l'image "hand".	129
4.4	Histogrammes lissés des composantes rouge, verte et bleue de l'image "hand".	130
4.5	Paramètres retournés par la méthode de détection des pics significatifs	132
4.6	Pixels de l'image "hand" appartenant aux sous-ensembles $S_i^{1,k}$ correspondant	
	aux pics significatifs des histogrammes des composantes $R, V$ et $B$	135
4.7	Histogrammes lissés de l'espace couleur $(L^*, u^*, v^*)$ pour l'image "hand"	136
4.8	Pixels de l'image "hand" dont les niveaux de couleur sont inclus dans les	
	intervalles délimitant les pics significatifs de l'histogramme lissé $h^{11,2}(x,\tau)$	
	de la composante $u^*$	137
4.9	Partition du plan couleur $(L^*, v^*)$ avec les seuils $T_{g,i}^{11,1}, T_{d,i}^{11,1}, T_{g,i}^{11,3}$ et $T_{d,i}^{11,3}$ .	139
4.10	Résultats de la première itération de l'algorithme de la figure 4.1 pour l'image	
	"hand"	140
4.11	Résultats de l'image "hand".	141
4.12	Segmentation de l'image "house" avec la méthode proposée	143
4.13	Segmentation de l'image "peppers" avec la méthode proposée	145
4.14	Segmentation de l'image "jeally beans" avec la méthode proposée	147
5.1	Poste d'acquisition.	152
5.2	Principe de fonctionnement du poste d'acquisition.	153
5.3	Courbes de sensibilité spectrale du capteur linéaire couleur Kodak KLI-2113	
	équipant la caméra Basler $L301 - kc$ (données constructeur)	154
5.4	Représentation spectrale relative d'énergie de la lampe à halogénure-métallique	
	utilisée (courbe constructeur)	155
5.5	Illustration d'un verre avec un défaut de "tâché"	157
5.6	Illustration d'un verre avec un défaut de "manque"	158
5.7	Illustration d'un verre avec un défaut de "décalage"	158
5.8	Illustration d'un verre avec un défaut de "nuance"	159
5.9	Algorithme de la procédure hors-ligne	162
5.10	Positionnement des ROI maîtres et esclaves et agrandissement de la ROI esclave	e164
5.11	Histogrammes des composantes couleur de l'espace $(Y, U, V)$ $(i = 6)$ à la pre-	
	mière itération appliquée à la ROI de la figure 5.10(b)	166
5.12	Résultats de la première itération de l'algorithme de la figure 5.9 appliquée à	
	la ROI de référence 5.10(b)	167
5.13	Résultats de la ROI "Fanta".	168

5.14	Algorithme de la procédure en-ligne	169
5.15	Comparaison des ROI.	172
5.16	Contrôle d'aspect d'une ROI esclave avec une seule couleur homogène ( $N_C =$	
	1) par réponse à un gradient couleur	174
5.17	Contrôle d'aspect d'une ROI esclave avec cinq classes de pixels ( $N_C = 5$ ) par	
	la méthode de multi-seuillage proposée.	175

# **Table des tableaux**

2.1	Valeurs des paramètres $\omega_{\min}$ , $\omega_{\min}$ et $\omega_{\max}$ des espaces triangulaires, hexago-	
	nal simple et hexagonal double à partir du modèle général GLHS.	68
2.2	Segmentation et espace couleur	83
4.1	Degré de connexité des sous-ensembles de pixels $S_i^{1,k}$ dont les niveaux de	
	couleur appartiennent aux intervalles $\left T_{g,i}^{1,k}, T_{d,i}^{1,k}\right $ pour l'image "hand"	133
4.2	Pouvoir discriminant $R^j$ de l'espace couleur $j$ et $R^{j,k}$ de ses composantes me-	
	surés à partir de l'image "hand" lors de la première itération.	136
4.3	Pouvoir discriminant des sous-ensembles de pixels $S_i^{11,k}$ dont les niveaux de	
	couleur appartiennent aux intervalles $[T_{g,i}^{11,k}, T_{d,i}^{11,k}]$ pour l'image "hand" codé	
	dans le système $(L^*, u^*, v^*)$	137
4.4	Espace couleur J sélectionné à chaque itération, pouvoir discriminant $R^J$ et	
	nombre $N_P^{J,K}$ de pics de la composante K la plus discriminante pour chaque	
	itération de la procédure appliquée à l'image "house"	143
4.5	Espace couleur J sélectionné à chaque itération, pouvoir discriminant $R^J$ et	
	nombre $N_P^{J,K}$ de pics de la composante K la plus discriminante pour chaque	
	itération de la procédure appliquée à l'image "peppers"	145
4.6	Espace couleur J sélectionné à chaque itération, pouvoir discriminant $R^J$ et	
	nombre $N_P^{J,K}$ de pics de la composante K la plus discriminante pour chaque	
	itération de la procédure appliquée à l'image "jelly beans"	147
5.1	Soustraction de deux images.	160
5.2	Calcul des pouvoirs discriminants à la première itération appliquée à la ROI	
	de la figure 5.10(b)	166
5.3	Résultats obtenus par notre méthode appliquée aux images	175

#### RÉSUMÉ en français

Les méthodes de segmentation analysant la distribution des couleurs dans un espace couleur supposent que les pixels appartenant à chaque région de l'image donnent naissance à un nuage de points dans un espace couleur. La segmentation consiste à identifier chaque nuage de points afin de construire des classes de pixels par des méthodes classiques d'analyse de données.

La couleur d'un pixel peut-être représentée dans différents espaces couleur qui respectent des propriétés spécifiques. La problématique que nous abordons est la sélection d'espaces couleur les mieux adaptés pour construire les classes de pixels. Nous proposons de nous placer dans un contexte non supervisé pour comparer le pouvoir de discrimination des différents espaces couleur afin d'en sélectionner le mieux adapté à la construction de chaque classe de pixels. La méthodologie proposée s'appuie sur une procédure itérative originale analysant conjointement les histogrammes monodimensionnels des composantes couleur associées aux espaces comparés et la répartition spatiale des pixels dans l'image. A chaque itération, une classe de pixels est extraite de l'image par multi-seuillage des histogrammes dans l'espace couleur sélectionné.

Notre méthodologie a montré des résultats très encourageants dans le cadre d'une application industrielle cherchant à détecter des défauts d'aspect apparaissant à la surface de verres ornés de motifs colorés. En raison de la forme cylindrique des verres, les images des surfaces des verres sont acquises par une caméra linéaire couleur. Un soin particulier a été apporté au développement du poste d'acquisition des images afin de mettre en évidence les défauts dans les images acquises.

## **TITRE** en anglais

Automatic selection of color spaces for color image segmentation - Application to the quality control of decorated glasses printing by a silk-screen process

## **RÉSUMÉ** en anglais

The color image segmentation techniques which analyze the distribution of the pixel colors in a color space assume that the colors of the pixels which belong to the regions of the image give rise to clusters in the color space. The segmentation consists to determine each cluster in order to build pixel classes by the use of the classical multi-dimensional data analysis techniques.

The pixel colors can be coded in several color spaces with specific properties. The key problem is the selection of relevant color spaces in order to build the pixels classes of the analyzed image. We propose to work with an unsupervised context in order to select an adapted color space for the construction of each class by the comparison of their discriminating power. This proposed method is based on an original iterative procedure of the mono-dimensional histograms associated with the color components of the compared color spaces. This method takes into account both the color distribution of the pixels in the selected color space and the spatial arrangement of the colors within the image. At each iteration step, one class of pixels is extracted by multi-thresholding in the selected color space.

This approach is applied for an industrial application and has provided very interesting results for the aspect flaw detection of decorated drinking glasses. Because of the cylindrical shape of the glasses, the images of the surfaces have been grabbed by a color line scan camera. The acquisition system development has been carefully studied in order to underline the aspect flaw in the grabbed images.

#### DISCIPLINE

Automatique et Informatique Industrielle.

#### **MOTS-CLÉS**

image couleur, segmentation, classification, espaces couleur, contrôle d'aspect.

Laboratoire d'Automatique **LAGIS**. Bâtiment P2 – Université des Sciences et Technologies de Lille – 59655 Villeneuve d'Ascq