



**HAL**  
open science

# Segmentation d'image: recherche d'une mise en oeuvre automatique par coopération de méthodes

Chafik Kermad

► **To cite this version:**

Chafik Kermad. Segmentation d'image: recherche d'une mise en oeuvre automatique par coopération de méthodes. Traitement du signal et de l'image [eess.SP]. Université Rennes 1, 1997. Français. NNT: . tel-00008781

**HAL Id: tel-00008781**

**<https://theses.hal.science/tel-00008781>**

Submitted on 18 Mar 2005

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

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

# Table des matières

<b>Introduction Générale</b>	<b>1</b>
<b>1 Concepts de la vision humaine et de la vision artificielle</b>	<b>5</b>
1.1 <b>Quelques mécanismes de fonctionnement du système visuel humain</b>	6
1.1.1 Introduction . . . . .	6
1.1.2 Perception visuelle . . . . .	6
1.1.2.1 Perception du contraste et localisation des objets dans une scène . . . . .	6
1.1.2.2 Routines visuelles . . . . .	8
1.1.2.3 Discrimination de textures . . . . .	8
1.2 <b>Vision Artificielle</b> . . . . .	9
1.2.1 Introduction . . . . .	9
1.2.2 Différentes approches de la vision artificielle . . . . .	9
1.2.2.1 Approche classique . . . . .	10
1.2.2.2 Vision active et vision animée . . . . .	10
1.2.2.3 Vision intentionnelle et vision qualitative . . . . .	11
1.2.2.4 Perception active . . . . .	12
1.3 <b>Conclusion</b> . . . . .	13
<b>2 Segmentation d'images</b>	<b>15</b>
2.1 <b>Introduction</b> . . . . .	15
2.2 <b>Approches non coopératives</b> . . . . .	16
2.2.1 Extraction de régions non texturées . . . . .	16
2.2.1.1 Introduction . . . . .	16
2.2.1.2 Méthodes de segmentation par seuillage . . . . .	17
2.2.1.2.1 Introduction . . . . .	17
2.2.1.2.2 Approches globales . . . . .	18
2.2.1.2.3 Approches locales . . . . .	29
2.2.1.2.4 Evaluation des techniques de seuillage . . . . .	31
2.2.1.2.5 Conclusions sur les techniques de seuillage . . . . .	33
2.2.1.3 Méthodes de segmentation par croissance de régions . . . . .	33
2.2.1.3.1 Introduction . . . . .	33
2.2.1.3.2 Approches par fusion . . . . .	35

	2.2.1.3.3	Approches par division . . . . .	37
	2.2.1.3.4	Approches par division-fusion . . . . .	37
	2.2.1.4	Conclusion . . . . .	39
2.2.2		Détection de contours . . . . .	40
	2.2.2.1	Introduction . . . . .	40
	2.2.2.2	Opérateurs du premier et du second ordre . . . . .	41
	2.2.2.3	Approches utilisant une modélisation surfacique . . . . .	41
	2.2.2.4	Approches par optimisation . . . . .	42
	2.2.2.4.1	Filtre de Canny . . . . .	42
	2.2.2.4.2	Filtre de Deriche . . . . .	44
	2.2.2.4.3	Filtre de Shen et Castan . . . . .	45
	2.2.2.4.4	Conclusion . . . . .	46
	2.2.2.5	Autres approches de détection de contours . . . . .	47
	2.2.2.6	Quelques méthodes de fermeture de contours . . . . .	48
	2.2.2.7	Conclusion . . . . .	50
2.2.3		Extraction de régions texturées . . . . .	51
	2.2.3.1	Définition et présentation du problème . . . . .	51
	2.2.3.2	Approches par extraction de paramètres et classification . . . . .	53
	2.2.3.2.1	Histogramme généralisé des niveaux de gris . . . . .	54
	2.2.3.2.2	Matrices de co-occurrences . . . . .	55
	2.2.3.2.3	Matrices de longueurs de plages . . . . .	56
	2.2.3.2.4	Analyse de texture par filtrage . . . . .	58
	2.2.3.2.5	Paramètres géométriques . . . . .	60
	2.2.3.3	Approche par modélisation de texture . . . . .	60
	2.2.3.3.1	Modélisation linéaire de la texture . . . . .	61
	2.2.3.3.2	Modélisation par champs de Markov . . . . .	63
	2.2.3.4	Conclusion . . . . .	67
2.3		<b>Approches coopératives . . . . .</b>	68
	2.3.1	Introduction . . . . .	68
	2.3.2	Approches séquentielles . . . . .	69
	2.3.2.1	Contrôle de la fusion des régions par les contours . . . . .	69
	2.3.2.2	Coopération contours-modèle surfacique . . . . .	69
	2.3.2.3	Coopération région-région et région-contour . . . . .	70
	2.3.2.4	Coopération division-fusion et contours actifs . . . . .	70
	2.3.2.5	Coopération squelettes-contour-région . . . . .	70
	2.3.2.6	Coopération contour-région fondée sur le diagramme de Voronoï . . . . .	71
2.3.3		Approches parallèles . . . . .	71
	2.3.3.1	Intégration de l'extraction de région et de contour . . . . .	71
	2.3.3.2	Extraction des régions et des contours par matrice de co-occurrence . . . . .	72
2.3.4		Approches hybrides . . . . .	72
	2.3.4.1	Fusion de plusieurs segmentations . . . . .	72

2.3.4.2	Coopération contour-région utilisant un système expert . . .	74
2.3.4.3	Coopération contour-région par approche multi-agents . . .	74
2.3.4.4	Extraction d'une segmentation consensus . . . . .	75
2.3.4.5	Coopération dynamique en boucle fermée . . . . .	76
2.4	<b>Méthodes d'évaluation des résultats de segmentation</b> . . . . .	78
2.4.1	Evaluation avec connaissance de la vérité de terrain . . . . .	78
2.4.1.1	Mesure de la performance de la localisation . . . . .	79
2.4.1.2	Mesure de la performance d'une détection . . . . .	80
2.4.2	Evaluation sans connaissance de la vérité de terrain . . . . .	80
2.4.3	Méthodes analytiques d'évaluation . . . . .	82
2.5	<b>Conclusion</b> . . . . .	82
<b>3</b>	<b>Système de segmentation proposé</b> . . . . .	<b>85</b>
3.1	<b>Introduction et principe général</b> . . . . .	85
3.2	<b>Extraction de régions uniformes et/ou faiblement texturées</b> . . . . .	86
3.2.1	Approche adoptée pour la détection des contours . . . . .	87
3.2.2	Approche adoptée pour l'extraction de régions uniformes . . . . .	90
3.2.2.1	Présentation des algorithmes de croissance de régions . . . . .	90
3.2.2.1.1	Algorithme idéal . . . . .	91
3.2.2.1.2	Algorithme récursif . . . . .	91
3.2.2.1.3	Algorithme récursif optimal . . . . .	92
3.2.2.1.4	Algorithme isotrope . . . . .	92
3.2.2.2	Choix d'un algorithme de croissance de régions . . . . .	93
3.2.2.3	Critère de similarité . . . . .	94
3.2.2.4	Détermination des régions-noyaux . . . . .	94
3.2.2.5	Modèles et structures des données . . . . .	95
3.2.2.5.1	Initialisation et mise à jour de la région . . . . .	95
3.2.2.5.2	Estimation récursive des paramètres de la région . . . . .	96
3.2.2.6	Résultats expérimentaux et conclusions . . . . .	96
3.2.3	Coopération contours-régions uniformes . . . . .	98
3.2.3.1	Interaction contours-régions uniformes . . . . .	100
3.2.3.1.1	Croissance de régions avec contrôle des contours . . . . .	100
3.2.3.1.2	Affinement et correction des contours en utilisant les régions . . . . .	101
3.2.3.2	Mesure de la correspondance contours-régions . . . . .	105
3.2.3.2.1	Position du problème . . . . .	105
3.2.3.2.2	Mesure de similarité entre contours et régions . . . . .	107
3.2.4	Résultats expérimentaux . . . . .	111
3.3	<b>Intégration de l'information texture dans la coopération contours- régions</b> . . . . .	<b>118</b>
3.3.1	Localisation des régions texturées . . . . .	119
3.3.1.1	Simplification par multi-seuillage . . . . .	119

3.3.1.1.1	Construction de l'histogramme global . . . . .	120
3.3.1.1.2	Classification des niveaux de gris . . . . .	125
3.3.1.1.3	Résultats expérimentaux . . . . .	127
3.3.1.1.4	Etudes comparatives et évaluation des méthodes de multi-seuillage . . . . .	131
3.3.1.1.5	Conclusion . . . . .	144
3.3.1.2	Extraction des textures par caractérisation et fusion des ré- gions multi-seuillées . . . . .	144
3.3.1.2.1	Introduction . . . . .	144
3.3.1.2.2	Localisation grossière des textures . . . . .	145
3.3.1.2.3	Affinement des frontières par modélisation auto- régressive bidimensionnelle de la texture . . . . .	150
3.3.1.2.4	Conclusion . . . . .	153
3.3.2	Interaction contour-régions uniformes et régions texturées . . . . .	154
3.3.3	Résultats expérimentaux . . . . .	155
3.3.4	Evaluation et étude comparative . . . . .	157
3.4	<b>Conclusion</b> . . . . .	168
	 <b>Conclusion et Perspectives</b>	 <b>171</b>
	 <b>Annexes</b>	 <b>174</b>
	<b>A Implémentation récursive de l'opérateur de Deriche</b>	<b>175</b>
	A.1 Cas monodimensionnel . . . . .	175
	A.2 Cas bidimensionnel . . . . .	178
	<b>B Lissage adaptatif utilisant la diffusion anisotropique</b>	<b>181</b>
	 <b>Bibliographie</b>	 <b>184</b>
	<b>Liste des tableaux</b>	<b>201</b>
	<b>Liste des figures</b>	<b>203</b>



# Introduction Générale

L'information visuelle est sans doute la plus riche des différentes sources d'informations disponibles. De ce fait, la conception des systèmes de vision pour l'interprétation automatique de scènes suscite un intérêt sans cesse croissant.

De nombreux chercheurs se sont penchés sur ce problème et ont mis en œuvre plusieurs approches en vue de la conception d'un système de vision complet. L'objectif des premières étapes de la plupart des systèmes de vision est de décrire l'importante quantité d'informations contenues dans l'image en recherchant des indices visuels ou des primitives pertinents permettant de la représenter sous une forme plus condensée et facilement exploitable. La performance des systèmes de vision artificielle est tributaire de la qualité de cette représentation.

Le cadre général dans lequel s'inscrit cette thèse est celui de la segmentation d'images. Cette opération constitue la base de tous les processus d'interprétation d'images et c'est à ce titre qu'elle reste encore l'un des principaux thèmes de recherche en vision artificielle.

Bien que le sujet ait déjà fait l'objet d'une vaste littérature [280] [289], il n'existe pas de méthode générale de segmentation [93]. La technique choisie reste généralement dépendante :

- des spécificités de l'image à traiter : (richesse en textures d'orientations et d'échelles différentes; non homogénéité de régions; transitions floues entre régions; contours occultes; conditions d'acquisition; présence de dégradations (bruits, flou); conditions d'éclairage; présence de reflets et d'ombres;...);
- du type d'indices visuels à extraire : (contours; régions uniformes au sens des niveaux de gris; textures; points caractéristiques; formes : segments de droite, angles;...);
- de la nature du problème à résoudre en aval de la segmentation : (quantification; classification; reconstruction 3D; reconnaissance de formes; interprétation; contrôle de qualité; localisation; suivi automatique d'objets;...);
- de contraintes d'exploitation : (complexité algorithmique; fonctionnement en temps réel; contraintes matérielles liées aux systèmes de capture, capacité de stockage;...).

Le problème de segmentation reste ouvert, surtout dans le cadre d'analyse d'images réelles riches en détails et contenant une grande variété de régions de nature uniforme ou texturée. En effet, comme c'est le cas pour de nombreuses tâches de vision artificielle, la segmentation d'images demeure un problème mal posé au sens de Hadamard [324] à cause de la non unicité, de l'incertitude et de l'instabilité des solutions par rapport aux données d'entrée. L'application de différents algorithmes à une même image donne souvent des résultats différents, avec une redondance parfois très variable. De faibles modifications des données initiales et des paramètres de la méthode (seuils, facteurs d'échelles, tailles de fenêtres d'analyse, sens de parcours des points,...) peuvent impliquer des changements remarquables sur les résultats.

Dans le but d'obtenir une segmentation efficace et suffisamment corrélée avec les entités qui composent la scène réelle, de nombreux travaux utilisent plusieurs techniques et adaptent chacune d'elles à la nature de la zone de l'image à traiter (prise en compte de l'homogénéité, de la texture, des contrastes, du bruit,...). Pour l'analyse des images complexes, il nous semble nécessaire d'*adopter une démarche progressive où la formation des primitives s'opère de manière coopérative et guidée; et de mesurer le degré de cohérence des primitives extraites par les différentes techniques*. Ce type de mesure par comparaison entre les résultats issus de chacune des techniques impliquées dans la coopération, fait défaut à bon nombre de méthodes proposées dans la littérature.

Les travaux menés dans le cadre de cette thèse portent sur différents aspects :

- l'étude du problème de la détection des indices visuels et l'examen des principales méthodes de segmentation d'images afin de dresser un état de l'art dans ce domaine;
- l'élaboration et l'analyse des critères permettant l'évaluation quantitative de la qualité des résultats obtenus par un algorithme de segmentation donné.
- la recherche d'une architecture d'un système de segmentation automatique ou semi-automatique et adaptable à l'analyse de plusieurs types d'images.

L'architecture du système que nous proposons combine deux concepts.

Le premier est fondé sur un procédé d'intégration d'informations résultantes de segmentations multiples en vue de tirer parti des avantages de chacune d'elles. Dans le cadre de notre travail, nous nous intéressons plus particulièrement à la fusion des résultats de segmentation de type contours et régions.

Le deuxième concept s'inspire de la perception active par l'introduction d'une boucle de retour dans le système. D'une manière générale, l'intérêt d'un tel mécanisme est de permettre la remise en cause et la correction des résultats obtenus ainsi que l'ajustement des paramètres de contrôle des différentes techniques de segmentation.

Dans notre démarche, contrairement à la plupart des travaux en segmentation déjà menés sur le sujet, nous ne nous sommes pas arrêtés au niveau de la coopération. Nous

avons poussé l'investigation jusqu'à la mesure de la cohérence entre les résultats engendrés par les techniques qui coopèrent.

Indépendamment de ses annexes, ce mémoire est organisé en trois chapitres.

**Le premier chapitre est consacré aux concepts de la vision humaine et de la vision artificielle.** Nous commençons par une mise en évidence des concepts qui seront développés dans la suite de ce travail (*indices visuels, perception visuelle*). Une présentation générale des différentes *approches de vision artificielle* proposées dans la littérature est ensuite effectuée.

**Dans le deuxième chapitre, nous passons en revue les principales approches de segmentation d'images.** Ce chapitre comporte cinq sections. Dans **la première**, une présentation de la segmentation en régions est réalisée. Nous étudions avec plus d'attention la *segmentation dans l'espace paramétrique* (multi-seuillage) et la *segmentation par agrégation de points*. Dans **la deuxième** section, nous faisons un *inventaire des principales méthodes de détection de contours*. Nous détaillons plus explicitement *la détection par filtrage optimal* (approche qui sera retenue dans la mise en œuvre de notre système de segmentation). Dans **la troisième** section, nous présentons les méthodes statistiques de segmentation de textures en les classant en deux catégories : les méthodes basées sur l'extraction et la classification de paramètres; et celles fondées sur une modélisation (markovienne, autorégressive,...). Dans **la quatrième** section, nous présentons les approches multi-primitives (coopératives). Nous proposons une *classification des différentes méthodes* rencontrées dans la littérature en trois catégories (*coopération séquentielle, coopération parallèle et coopération hybride*). Nous étudions ensuite les trois schémas et présentons quelques exemples pour chacun des schémas. **La dernière** section aborde le problème de l'évaluation quantitative de la qualité des résultats obtenus par une méthode de segmentation donnée.

**Le troisième et dernier chapitre porte sur l'approche de segmentation proposée et donne quelques résultats expérimentaux ainsi qu'une étude comparative.** Dans ce chapitre, nous présentons l'architecture générale du système coopératif et itératif développé en décrivant les deux modules qui le composent.

Le **premier** est dédié à l'extraction de régions uniformes ou faiblement texturées. Il est constitué de deux méthodes : détecteur optimal de *contours* et extracteur de *régions* par agrégation de points. Ces deux méthodes coopèrent mutuellement de façon itérative. A chaque itération, une mesure du degré de cohérence entre les résultats issus de chacune de ces deux méthodes est effectuée. La cohérence est réalisée en minimisant une mesure de dissimilarité entre les résultats. Le but est ainsi de fournir une solution optimale au sens de la compatibilité entre les deux segmentations.

Le **second module** fait intervenir les primitives "textures" afin de réactualiser et corriger les primitives contours et régions non significatives extraites au fur et à mesure des différentes itérations du premier module. Les primitives "textures" sont localisées à

l'aide d'un algorithme opérant suivant deux niveaux : division et fusion. La division est réalisée à l'aide d'une technique de multi-seuillage utilisant des critères d'homogénéité locale prenant en compte des propriétés du système visuel humain liées à la sensibilité de l'oeil aux contrastes. La fusion est basée sur les co-occurrences des niveaux de gris issus de la division.

Pour chacun des modules du système, nous détaillons les méthodes développées et nous donnons des résultats comparatifs sur des images synthétiques et réelles issues de scènes médicales, satellitaires, aériennes et d'extérieur. Nous étudions, ensuite, l'interaction entre les deux modules, puis, nous présentons des résultats obtenus sur une image aérienne riche en détails et contenant une grande variété de régions de nature uniforme et texturée. Enfin, une étude comparative avec des méthodes développées par ailleurs, sur une image satellitaire pour laquelle nous disposons d'une vérité de terrain, permet de déterminer les forces et faiblesses de notre approche.

**En conclusion générale**, nous résumerons notre contribution et tirerons les conclusions sur l'approche développée. Nous proposerons, enfin, quelques unes des améliorations qui pourraient être apportées.

# Chapitre 1

## Concepts de la vision humaine et de la vision artificielle

Dans ce chapitre, nous rappelons quelques notions liées à la vision humaine et plus particulièrement celles ayant un lien avec l'extraction d'indices visuels. Ensuite, nous exposons les différentes approches actuelles de vision artificielle et discutons de la méthodologie qui a inspiré nos travaux, c'est-à-dire la perception active.

Nous pouvons considérer la vision humaine comme un système général au niveau de ses mécanismes de compréhension et de raisonnement appropriés. Ces mécanismes permettent l'interprétation instantanée de scènes très complexes et variées. Ils font appel à des connaissances a priori de ce que nous observons, comme les modèles et les catégorisations des objets, et à la maîtrise des stratégies de commande de nos capteurs visuels, comme la fixation, la focalisation et le suivi d'objets. Le système visuel humain utilise, pour détecter les objets, plusieurs informations, notamment les intensités lumineuses réfléchies par les objets présents, les différences de niveaux de gris ou de textures entre ces objets, etc. Doté de l'ensemble des éléments précédents, le système visuel humain exécute de manière triviale et immédiate les tâches de segmentation d'images et de reconnaissance de formes. En revanche, ces tâches posent de nombreuses difficultés non résolues en vision artificielle. La raison de cet échec est en partie due à une utilisation trop limitée de l'information contenue dans l'image ainsi qu'à une modélisation trop simpliste de cette information [100] [259] [334].

Notre compréhension de la vision humaine reste partielle: il n'existe pas encore, à notre connaissance, de théorie globale et précise dans ce domaine. Il y a quelques modèles quantitatifs qui tentent de décrire la phase préattentive de la vision mais dès qu'il s'agit d'aborder des phénomènes plus cognitifs ou le processus de vision dans son ensemble, on manie plutôt des modèles qualitatifs, trop vagues pour une implantation sur ordinateur [90] [222] [314] [348].

Les paragraphes suivants sont consacrés à une présentation succincte d'un certain nombre d'éléments relatifs à la perception visuelle humaine et à la vision artificielle.

## 1.1 Quelques mécanismes de fonctionnement du système visuel humain

### 1.1.1 Introduction

Les travaux sur l'analyse du système visuel humain ont permis de comprendre relativement bien les premières étapes de son fonctionnement notamment grâce aux études psychophysiologiques et neurobiologiques. Si beaucoup d'inconnues subsistent dans les détails (le fonctionnement d'un simple neurone est encore loin d'être complètement analysé), l'étude globale du cerveau montre qu'il existe des régions où les neurones sont différemment organisés. Ils sont souvent regroupés en couche avec des spécificités propres, tant au niveau des connexions avec les autres couches qu'au niveau des propriétés internes comme les seuils d'excitabilité et de sensibilité. Ces éléments permettent de supposer qu'il existe des groupes de neurones dédiés à une tâche particulière. Les différentes tâches sont généralement exécutées de manière parallèle. La vision pourrait ainsi être décomposée en **plusieurs étapes distinctes**. Des expériences ont par exemple montré que certaines zones sont sensibles à la couleur alors que d'autres sont sensibles à la forme.

### 1.1.2 Perception visuelle

#### 1.1.2.1 Perception du contraste et localisation des objets dans une scène

De nombreux travaux ont été effectués dans le but de mettre en œuvre une théorie de *la perception visuelle* chez l'homme. La psychophysique, théorie fondée par Weber et Fechner [90] [148], propose d'étudier l'ensemble des phénomènes liés à la perception. Elle consiste à examiner quantitativement les relations qui relient l'intensité physique d'une stimulation sensorielle à l'intensité subjective de la sensation qui en résulte. Le seuil différentiel est la quantité minimale dont il faut accroître un stimulus pour qu'un changement soit perçu. Weber a tenté de montrer que le seuil différentiel est une fraction constante du stimulus, quelle que soit l'intensité de ce dernier.

Dans une seconde étape, Fechner émit le postulat d'après lequel l'échelon différentiel serait l'unité de mesure de la sensation. En outre, en supposant la constance du rapport  $(\Delta B/B)$  représentant le contraste, mis en évidence par Weber pour les seuils différentiels ( $B$  étant l'intensité du stimulus), Fechner considère que l'intensité de la sensation est proportionnelle au logarithme de l'intensité du stimulus.

En ce qui concerne les sensations de brillance, le seuil différentiel pour la brillance ne s'exprime pas toujours par une fraction constante du stimulus lumineux.  $\Delta B/B$  n'est constant que pour des intensités moyennes de luminance mais augmente par contre fortement lorsque l'intensité de la luminance est très petite ou très grande. La formule de Fechner ne peut donc pas être appliquée ici, étant donné que la loi de Weber ne se vérifie pas sur le plan expérimental.

Stevens a repris l'ensemble du problème de la psychophysique et a proposé deux méthodes différentes permettant d'établir une échelle de sensations [148]. Il a constaté au moyen de ces deux méthodes, que des stimulations sensorielles additives, telles que les échelles d'intensité sonore ou d'intensité lumineuse, obéissent à une loi de puissance  $Y = kB^n$ . Par contre, lorsqu'on utilise des stimulations sensorielles impliquant des changements du siège de l'excitant, telles qu'une échelle de hauteur tonale, c'est la loi de Fechner qui se vérifie expérimentalement, même si on utilise la méthode d'estimation directe.

Dans une autre étude proposée par Liao et Chehdi [210], une fonction de sensibilité au contraste a été établie. Pour mesurer la sensibilité de l'oeil au contraste, une série d'images composées de bandes verticales est créée. Chaque bande dans l'image correspond à un niveau de luminance constant. Partant du niveau de luminance faible (0) vers le niveau élevé (255), deux bandes successives sont nuancées par un contraste  $C$ . Pour un ensemble de valeurs de contraste, une série d'images est obtenue. L'observation de l'ensemble de cette série montre que trois zones sont généralement distinguées. Les zones extrêmes, correspondant aux bandes de luminances faible et élevée, sont pratiquement homogènes, c'est à dire que les frontières entre bandes d'une zone sont invisibles par un observateur humain. Par contre les bandes de la zone médiane se distinguent facilement et se propagent horizontalement du centre vers les deux côtés de l'image quand la valeur de  $C$  augmente. La perception du contraste entre un point et son voisinage est dépendante de son niveau de luminance, et de la différence des luminances de ses voisins proches.

L'oeil est moins sensible aux transitions dans les luminances élevées (zones très claires de l'image) et les luminances faibles (zones foncées). La fonction de perception visuelle établie a été utilisée dans une procédure de détection de contours afin de les répertorier en plusieurs catégories[210].

Concernant *la localisation des objets dans une scène*, A. Treisman [328] a élaboré une autre théorie de la perception visuelle chez l'être humain. La théorie s'applique exclusivement à l'exploration de phénomènes bas niveaux, son champ d'investigation ne couvre donc pas, par exemple, les processus du niveau cognitif. Elle fut initialement proposée pour expliquer les performances mesurées lors de diverses expériences sur la localisation des objets dans une scène. Dans sa version première, sont clairement séparées deux étapes du traitement de l'information visuelle. Lors de la première étape, l'étape préattentive, plusieurs cartes rétino-topiques de primitives sont dressées. Ces primitives sont par exemple, la texture, la couleur, l'orientation des segments ou la direction du mouvement. Elles sont extraites en parallèle via des canaux différents. Les discontinuités des cartes de primitives sont alors détectées et synthétisées dans une carte des positions qui donne l'emplacement des discontinuités mais aucune information sur l'identité de ou des cartes de primitives qui leur ont donné naissance. La deuxième étape focalise un faisceau attentif sur les discontinuités présentes dans la carte des positions. Chaque fixation a pour but la recherche des cartes de primitives responsables de la formation de la discontinuité observée.

### 1.1.2.2 Routines visuelles

Selon S. Ullman [338], une scène est appréhendée par le cortex visuel humain après déclenchement de séquences d'opérations élémentaires opérant sur une représentation primaire de la scène et spécialisées chacune dans l'extraction d'une caractéristique radiométrique, géométrique ou topologique bien localisée. La représentation de base, est obtenue par extraction de primitives locales sur l'image rétinienne. Les extractions locales se font de manière parallèle sur l'ensemble du champ visuel. Les primitives déterminées sont en fait inspirées de celles suggérées par D. Marr [222] dans son analyse de la vision: contours, couleurs, profondeurs, direction du mouvement,... La représentation de base est complètement indépendante de l'activité dans laquelle le cerveau est engagé (localiser tel objet dans une pièce); ainsi, pour une scène donnée, on obtiendra toujours la même représentation de base. La variété des formes possibles et des relations spatiales entre objets interdit le recours à un processus du type appariement modèle-observation pour l'identification d'une propriété géométrique ou d'un arrangement entre objets de la scène (le cardinal de l'ensemble des configurations possibles est infini!). S. Ullman suggère plutôt l'utilisation de séquences d'opérations élémentaires, chaque séquence définissant une routine visuelle. Ces opérations élémentaires sont pour beaucoup inspirées de la vision artificielle:

- **indexation:** des points remarquables sont recherchés dans la représentation de base pour fixer l'attention;
- **activation bornée ou coloriage:** un point du champ visuel est déclaré actif (il a par exemple été désigné par un processus cognitif de haut niveau), le mécanisme de coloriage propage alors l'activité du point source aux points adjacents sans traverser une frontière; le processus est réitéré jusqu'à invariance. Le processus de coloriage serait sollicité dans la détermination de relations spatiales (un point est-il à l'intérieur ou à l'extérieur d'une courbe fermée?);
- **tracé de contours:** ce mécanisme est spécialisé dans le suivi des zones de discontinuité dans la représentation de base;
- **marquage:** les régions de l'image rétinienne déjà examinées sont marquées afin de détecter des contours fermés ou de compter les points d'intérêt.

L'ensemble de ces éléments a inspiré notre démarche pour l'extraction des zones homogènes au sens des niveaux de gris.

### 1.1.2.3 Discrimination de textures

Les textons ont été introduits par B. Julesz [179] pour dévoiler les mécanismes de discrimination entre textures visuelles par le cortex cérébral humain. La collection des textons aujourd'hui reconnus va de la couleur à la disparité binoculaire en passant par l'orientation des lignes droites, les terminaisons des contours, ou simplement la luminance. Si la définition du terme texton demeure floue, il a été vérifié expérimentalement qu'une région est perçue

comme une région texturée et donc caractérisable par des textons, si les motifs de base qui composent la texture respectent une contrainte de proximité: deux motifs adjacents sont astreints à appartenir à un même voisinage (désigné en tant que  $i$ -voisinage dans la théorie) dont l'étendue est de l'ordre de deux fois la taille du motif (le critère est ici exprimé pour un motif compact pour lequel la taille a une signification claire).

## 1.2 Vision Artificielle

### 1.2.1 Introduction

Le but de la vision artificielle est de doter les machines d'un système visuel leur permettant, à partir d'images d'effectuer des tâches complexes comme, par exemple, la photo-interprétation, la télé-surveillance, la navigation ou la manipulation d'objets par des robots. Parmi les autres applications dans ce domaine nous pouvons citer: l'analyse d'images aériennes ou satellitaires (surveillance des cultures ou des forêts, prévisions météorologiques,...); l'analyse d'images médicales; l'analyse de documents; le contrôle qualité;...

Un système de vision peut être décrit comme un ensemble de processus fortement liés opérant des transformations progressives des informations sensorielles. Ces transformations assurent le passage de l'image physique saisie à une description structurale puis sémantique, en passant par la segmentation et l'identification d'attributs caractéristiques de la scène et des objets qui la composent. D'une manière générale, deux éléments dans la définition d'un système de vision peuvent être retenus [152]:

- l'espace des représentations des données utilisées par le processus d'interprétation d'image.
- les connaissances qui définissent le sens d'une interprétation. Elles représentent l'ensemble des relations existant au sein de l'espace des représentations.

La première méthodologie pour la construction d'un système de vision artificielle a été proposée par Marr [222]. Pendant des années, l'approche de Marr a guidé les recherches en vision et a permis des avancées significatives. Cependant, les applications industrielles des systèmes ainsi développés sont restées très limitées, la plupart du temps à un monde bien déterminé et connu d'avance. Les difficultés rencontrées lors du développement de systèmes de vision ont poussé un certain nombre de chercheurs à aborder ce problème de différentes façons donnant naissance à plusieurs écoles au sein de la communauté de recherche en vision. Chacune de ces écoles définit une approche différente pour aborder le problème de la vision artificielle et pour construire des systèmes.

### 1.2.2 Différentes approches de la vision artificielle

Différentes méthodologies de la vision artificielle ont été proposées, parmi lesquelles on peut citer le reconstructionisme, la vision active, la théorie de "Gestalt", le "smart sensing",

la perception active, la vision animée,...

Dans les paragraphes suivants, après la présentation de la démarche de Marr, nous résumons les différentes approches actuelles de la vision artificielle.

### 1.2.2.1 Approche classique

Proposée par Marr [222], cette approche vise la reconstruction d'une description du monde 3D en terme de surfaces et d'objets et de leurs relations physiques et spatiales. La démarche adoptée par Marr pour établir les fondements de l'interprétation d'image, se résume à la recherche des réponses aux questions suivantes:

- quelles sont les causes des variations d'intensité du niveau de gris dans les images?
- comment classer ces variations d'intensité en fonction des causes qui leur ont donné naissance?
- comment extraire des informations sur la forme des objets à partir de l'observation de la distribution de ces variations d'intensité?
- tous les types de variation d'intensité observables, sont-ils nécessaires pour l'interprétation? Si ce n'est pas le cas, quelle est la classe de variations qui est suffisante pour l'interprétation d'une image?

Trois niveaux pour comprendre/construire un module de traitement visuel sont alors définis:

- l'étude des entités à calculer à partir des images, les raisons pour lesquelles il faut les calculer, et comment les calculer. Elle se traduit par la construction de modèles mathématiques de différents phénomènes observés dans les images et servant à leur compréhension.
- les représentations et les transformations. Il s'agit de la définition des représentations des données d'entrée et de sortie d'un module de vision et les algorithmes de transformation de ces données.
- l'implantation. C'est l'étape finale consistant à implanter les algorithmes et les structures de données de la manière la plus appropriée à une machine cible.

### 1.2.2.2 Vision active et vision animée

Partant du principe que l'observateur humain est un observateur actif, la vision active se veut exploratoire, tentant d'imiter la perception humaine. Aloimonos et al. [14] perçoivent la plupart des problèmes fondamentaux de la vision comme étant des problèmes réputés mal posés dans le cadre de l'approche de Marr. Leur complexité non surmontable provient de l'observation passive. Un observateur actif se déplaçant d'une façon connue et contrôlable

acquiert **plus d'informations**. Contraints par des mesures supplémentaires, ces problèmes mal posés deviennent bien posés, avec des solutions uniques et plus stables.

La vision active a pour objectif de construire des modules robustes pour l'extraction des primitives d'image. D'un point de vue méthodologique, la vision active, où les informations perçues sont utilisées au sein d'une boucle de rétroaction, tente surtout d'améliorer la qualité de la perception par rapport à l'approche passive classique, où l'on se restreint à observer, mesurer et interpréter les données issues du capteur. La vision active consiste à élaborer des stratégies de perceptions intelligentes, en contrôlant les paramètres du capteur (position, vitesse, mise au point, etc.). Elle peut être définie comme un processus d'acquisition "intelligent" des données afin de résoudre les problèmes soulevés dans la conception d'un système de vision par l'approche de Marr (sensibilité au bruit et faible précision). Dans cette approche, un certain nombre de problèmes restent toutefois à résoudre, comme:

- le contrôle des trajectoires d'un capteur par rapport aux scènes observées et aux informations à extraire des images,
- les méthodes d'apprentissage de ce contrôle par un système,
- la calibration du capteur et du système.

La vision animée est une des dérivés de la vision active. Elle a été proposée par Ballard [21]. En se basant sur l'étude biologique de l'œil humain, surtout sur le fait que la résolution autour de l'axe optique est supérieure à la résolution périphérique, et sur les études des mouvements intentionnels de l'œil durant une tâche visuelle, Ballard soulève la nécessité d'inclure de tels mécanismes dans un système de vision artificielle. Un système de vision animée est un système doté de mécanismes de fixation rapide inspirés du comportement visuel des systèmes biologiques. Ballard illustre cette approche par un exemple de détection d'un petit objet (une tasse à café) en ayant une connaissance a priori de sa position spatiale par rapport à un grand objet (sur la table). Une fois le grand objet détecté, le mécanisme de zoom permet de détecter le petit.

### 1.2.2.3 Vision intentionnelle et vision qualitative

Afin de simplifier son approche dans la vision active, Aloimonos [5] propose un paradigme de construction de systèmes de vision qu'il appelle vision intentionnelle. S'opposant toujours à l'approche traditionnelle à cause des difficultés rencontrées par celle-ci dans la construction de système de vision robuste, Aloimonos soulève le problème des buts d'un système de vision. Un système de vision ne doit pas obligatoirement construire entièrement son environnement s'il n'en a pas besoin. Il faut alors garder à l'esprit le but pour lequel un système de vision doit être conçu et s'éloigner de l'approche consistant à résoudre le problème général de la vision dont les éléments de base s'avèrent pour le moment très difficiles à réaliser et loin d'être surmontables. Pour lui, beaucoup de tâches visuelles complexes peuvent être réalisées d'une façon performante et robuste pour résoudre des problèmes de robotique industrielle.

L'idée de base consiste à découper une tâche visuelle en un ensemble de tâches simples et de connecter ces tâches d'une façon à répondre le mieux au but du système de vision. Dans cette perspective, les modules d'un tel système deviennent des modules reconnaissant des indices visuels simples et des situations, permettant d'effectuer les actions appropriées, et non pas des modules de traitement pour construire le monde à partir des images. La deuxième idée de base proposée par Aloimonos est la vision qualitative. Beaucoup de mesures dans un système visuel peuvent être qualitatives et ne nécessitent pas une grande précision. Qualitative signifie que les questions auxquelles un module de vision doit répondre sont posées d'une manière telle que les réponses peuvent par exemple être Oui/Non.

Les approches de la vision qualitative et de la vision intentionnelle permettent de construire des systèmes de vision robustes, réalisant des tâches relativement complexes. Cependant, beaucoup de questions restent à résoudre pour arriver à des systèmes dont le but est de comprendre leur environnement pour pouvoir réaliser leurs tâches. La vision qualitative peut être efficace pour un certain type de système visuel, comme la navigation des robots [245]. Elle peut aussi répondre à un certain type de raisonnement dans un système d'interprétation, palliant les imprécisions des mesures extraites des images.

#### 1.2.2.4 Perception active

Cette approche a été proposée par R. Bajcsy dans [19] comme étant une approche d'acquisition intelligente des données. Dans ce contexte, le terme 'active' signifie aussi bien la commande des capteurs que la modification des paramètres d'état d'un module de traitement. Ces deux actions sont guidées par les buts du système de vision. Bajcsy définit la perception active comme l'étude de la modélisation et des stratégies de la perception visuelle. La modélisation est celle des capteurs, des modules de traitement et de leur interaction. La stratégie est la recherche d'une séquence d'étapes de traitement qui peut minimiser une fonction de perte dans une recherche d'un maximum d'informations. La modélisation dans un système de perception active est constituée de deux types de modèles: les modèles locaux et les modèles globaux.

- les modèles locaux sont caractérisés, à tous les niveaux, par certains paramètres des modules de traitement, comme les paramètres des capteurs ou des algorithmes de traitement d'image. Ces modèles permettent la prédiction du comportement/résultats d'un module. Par exemple, pour un algorithme de segmentation par croissance de région, les paramètres définissant la taille du voisinage et le critère de similarité locale, permettent de prédire la sensibilité/robustesse d'une segmentation.
- les modèles globaux représentent l'interaction entre les différents modules. Ils déterminent la façon par laquelle les résultats intermédiaires des modules sont combinés. Un modèle global intègre les paramètres globaux du système ainsi que son état final.

L'idée fondamentale de ces modèles est l'introduction des boucles de retour dans le système et la recherche des données au fur et à mesure qu'elles deviennent nécessaires.

Bajcsy justifie l'utilisation de son approche dans des processus de vision dont le contrôle est aussi bien guidé par les données [9] [199] que par les modèles.

Le synoptique de perception active suggéré par Bajcsy incite à la définition explicite des paramètres des différents modules de traitement (capteurs et procédures de traitement d'image) et de leurs incertitudes (le modèle local). Ces deux mesures seront utilisées en rétroaction à travers des boucles de contrôle intelligent (le modèle global).

## 1.3 Conclusion

Dans ce chapitre, nous avons vu que l'approche neurobiologique met en évidence la décomposition en plusieurs étapes du processus de vision ainsi que l'importance de la détection des discontinuités locales de types variés, les zones homogènes en intensité et en texture...

La psychophysiologie, ou étude du comportement, a mis en évidence l'étape nécessaire de regroupement perceptuel préattentif de primitives simples, indépendamment de la connaissance symbolique.

Nous avons également présenté les différentes approches actuelles de la vision artificielle, et discuté brièvement les questions posées sur le problème d'extraction de primitives, à savoir:

- quelles informations faut-il extraire des images et comment?
- quand faut-il les extraire, sachant qu'il est irréaliste d'extraire toutes les informations nécessaires en une seule fois?

Dans le cas de l'approche classique, les modules de vision sont examinés indépendamment les uns des autres, ils se basent sur une connaissance exacte des entrées/sorties de chaque module.

La vision active est fiable dans son approche d'extraction robuste de certaines primitives d'image et la construction de systèmes de vision réactifs [96] [174]. Le mouvement d'un capteur peut être utilisé pour des buts de plus haut niveau, comme la focalisation sur une partie de la scène (afin d'obtenir plus de détails), ou pour la prise de plusieurs vues du même objet (afin de supprimer une occlusion ou une ambiguïté d'interprétation). Dans cette perspective, l'idée de la commande d'un capteur se ressemble à celle proposée dans la vision animée, sans pour autant être le seul mécanisme intelligent dans un système d'interprétation. Toutefois, la vision active et la vision animée ne peuvent pas être employées dans tous les domaines d'application. Elles se limitent aux domaines où les capteurs peuvent se déplacer (robotique).

La vision intentionnelle présente l'avantage d'une méthodologie permettant l'interaction dynamique entre différents modules en fonction d'un objectif à atteindre. Cependant, cette interaction nécessite des mécanismes de contrôle ayant des connaissances sur les modules et sur les meilleures décisions à prendre dans une situation donnée. Ces mécanismes doivent

être explicites dans un système de vision. La vision qualitative est appropriée pour effectuer des tâches qui nécessitent des informations qualitatives non précises de la scène.

La perception active semble être l'approche la plus générale. Elle ne se limite pas aux commandes des capteurs, mais s'étend aux contrôles de tous les modules de traitement. Le modèle local permet de contrôler et d'adapter les algorithmes de traitement d'image à une tâche donnée. Le modèle global supervise les interactions entre les modules. Celles-ci peuvent être exprimées par des contraintes ajustées en fonction des résultats obtenus par un module ou en fonction de la tâche qu'on lui demande de réaliser. De ce fait, un module peut être contrôlé dans des phases de traitement dirigé par les données ou par l'objectif à atteindre.

Malgré cette diversité de points de vue, la majorité des chercheurs en vision [100] admettent les difficultés de l'approche traditionnelle et la nécessité de chercher des alternatives sans pour autant remettre en cause la méthodologie de Marr. La plupart soulignent que le problème actuel à étudier en vision est le contrôle intelligent de la sélection des tâches visuelles appropriées à chaque type d'indice visuel [48]. De plus, ils soulèvent la nécessité de toutes sortes de contraintes pour résoudre le problème de la vision. Ils proposent de focaliser la recherche sur les représentations les mieux adaptées et les plus pertinentes à chaque niveau du processus visuel, sur les liens entre perception et action d'une manière générale ainsi que sur les liens entre la perception et la résolution des problèmes [334].

Un système d'interprétation doit éventuellement faire coopérer plusieurs des différentes approches de la vision artificielle pour améliorer ses performances et atteindre son but. Il doit être capable de choisir et adapter le traitement d'extraction des primitives.

# Chapitre 2

## Segmentation d'images

### 2.1 Introduction

Le processus d'analyse d'images peut être défini comme l'ensemble de méthodes et d'outils permettant de décrire quantitativement le contenu d'une image. Il est généralement décomposé en plusieurs étapes: acquisition, numérisation, pré-traitement, segmentation, interprétation.

L'étape de segmentation d'image est une étape très importante dans cette chaîne d'analyse car c'est à partir de l'image segmentée que les mesures sont effectuées pour l'extraction des paramètres discriminants en vue de la classification ou de l'interprétation. La tâche de segmentation peut se résumer de la manière suivante:

*Etant donnée une image, l'objectif de la segmentation est d'établir une description compacte et représentative de son contenu informationnel, plus exploitable que l'ensemble de ses points. Il s'agit de procéder à l'extraction d'indices visuels (primitives) pertinents, suffisamment corrélés avec les entités qui composent la scène d'où l'image est prise.*

La diversité des travaux menés dans le but de la compréhension de la vision humaine montre la complexité de la notion d'indice visuel et qu'il est bien difficile d'en donner une définition précise. Ceci étant, un indice visuel peut être défini comme une information perceptible directement à partir de la visualisation de l'image (contours, régions homogènes au sens des niveaux de gris et au sens de la texture, formes élémentaires). Ces indices visuels ou primitives correspondent à des phénomènes photométriques et/ou géométriques remarquables.

Il existe dans la littérature une très grande diversité de techniques de segmentation que l'on peut classer en deux principales catégories [186]:

**les approches non coopératives et les approches coopératives.**

- Dans la catégorie des **approches non coopératives** on peut distinguer trois types de méthodes [167] [280] [289]:
  - *les méthodes adaptées à l'extraction des régions homogènes au sens des niveaux de gris.* La segmentation dans l'espace paramétrique (multi-seuillage ou clustering) entre dans ce cadre. On y adjoint aussi les méthodes par fusion et/ou division de régions ainsi que les méthodes par agrégation de points.
  - *les méthodes pour la détection de contours.* Entrent dans ce cadre les méthodes dérivatives, les méthodes surfaciques et variationnelles ainsi que les méthodes markoviennes. On y inclut les méthodes morphologiques.
  - *les méthodes s'appliquant à la localisation des zones homogènes au sens de la texture.* Dans ce cadre, s'inscrivent les méthodes basées sur l'analyse statistique (extraction de paramètres de co-occurrence, longueur de plages,...; modélisation markovienne, AR,...) ainsi que les méthodes structurelles (théorie de langage, graphes,...).
- **Les approches coopératives** combinent plusieurs des méthodes de la première catégorie, en particulier celles de détection de contours et d'extraction de régions uniformes. Dans cette catégorie, trois types de coopération peuvent ressortir: la *coopération "série"*, la *coopération "parallèle"*, et la *coopération "hybride"*.

La suite de ce chapitre traitent des différentes méthodes de ces deux catégories.

## 2.2 Approches non coopératives

### 2.2.1 Extraction de régions non texturées

#### 2.2.1.1 Introduction

L'objectif de cette section est de présenter la segmentation en régions non ou faiblement texturées. Elle comporte deux sous-sections. Nous étudierons, dans la première, la segmentation dans l'espace paramétrique (clustering ou multi-seuillage). Dans la seconde partie, la segmentation par fusion et division de régions ainsi que la segmentation par agrégation de points sont présentées.

Les méthodes de Segmentation de Régions Non Texturées (SRNT) consistent à rechercher des ensembles de points qui partagent des propriétés similaires. Les régions correspondent aux parties connexes de ces ensembles. Deux techniques de SRNT peuvent être distinguées:

- **segmentation par classification** : elle consiste à regrouper les pixels de niveaux semblables, indépendamment des relations de connexité qui peuvent les lier. La technique de seuillage d'histogrammes, qui constitue la majorité des méthodes de segmentation

par classification, s'appuie sur l'hypothèse que les régions de niveau de gris uniforme produisent des modes suffisamment significatifs dans les histogrammes de l'image pour que l'on puisse les caractériser directement par les valeurs limites des pixels qui les composent. Il suffit alors de seuiller l'image entre ces deux limites pour en extraire les régions. L'inconvénient de ces méthodes est qu'elles ne prennent pas en compte la topologie de l'image.

- **segmentation par croissance de régions** : elle permet de sélectionner un pixel ou un ensemble de pixels de l'image, appelé noyau, autour duquel on fait croître une région. Ce pixel est généralement choisi de manière optimale au sens d'un critère qui peut nécessiter l'utilisation des résultats de segmentation antérieurs. Le problème du choix du noyau peut être évité en utilisant un algorithme de type division/fusion. Dans ce cas, l'image est divisée en petites zones de niveau uniforme. Les zones adjacentes sont alors fusionnées en des régions si elles satisfont un critère de similitude. Ces méthodes souffrent néanmoins du choix du critère d'arrêt de la croissance.

Nous décrivons dans les paragraphes suivants quelques unes des méthodes de chacune de ces deux approches. Nous discutons les limites et les avantages de chacune d'elles afin d'intégrer éventuellement les meilleures d'entre elles dans l'approche coopérative de segmentation que nous proposons.

### 2.2.1.2 Méthodes de segmentation par seuillage

#### 2.2.1.2.1 Introduction

La notion d'image qui est utilisée dans la suite est de nature bidimensionnelle discrète. A chaque élément (point de coordonnées  $(i, j)$ ) de l'image  $I$  composée de  $M \times N$  éléments, correspond un niveau d'intensité lumineuse  $I(i, j)$  (appelé niveau de gris) appartenant à l'ensemble  $G = \{g_0, g_1, \dots, g_{N_G-1}\}$  où  $N_G$  correspond au nombre total des niveaux de gris.

On définit l'histogramme  $h$  des niveaux de gris d'une image comme étant la fonction qui associe à chaque niveau de gris  $g_i$  ( $i$  compris entre 0 et  $N_G - 1$ ) le nombre de pixels de l'image  $h(g_i)$  qui possèdent cette intensité lumineuse.

La segmentation par seuillage est l'opération qui associe à chaque point de l'image une classe de luminance après comparaison de ses attributs (niveaux de gris) à un ou plusieurs seuils. Une étiquette  $e$  est alors affectée à chacun des points en fonction de la classe d'appartenance.

Si  $p_{ij}$  dénote un point de l'image de coordonnées  $(i, j)$ , l'opération de seuillage peut être formalisée par la relation suivante:

$$C_{p_{ij}} = e \quad \text{si} \quad g_{l-1} \leq A_{p_{ij}} < g_l, \quad l = 1, 2, \dots, m. \quad (2.1)$$

où  $C_{p_{ij}}$  correspond au résultat de la classification du point  $p_{ij}$ ,  
 $e$  est une étiquette.  
 $g_{l-1}$  et  $g_l$  représentent les valeurs des seuils (niveaux de gris).  
 $A_{p_{ij}}$  est une fonction caractéristique du point  $p_{ij}$  (par exemple  $I(i, j)$ )  
 $m$  correspond au nombre de classes associé à l'image segmentée.

D'une façon générale, la classification des points de l'image peut s'effectuer de deux manières différentes. Dans le premier cas, les seuils sont déterminés *globalement* pour tous les points de l'image, alors que dans le second cas le calcul des seuils est fait *localement* et de manière adaptative à l'ensemble des points auxquels ils vont être appliqués.

Nous allons passer en revue les méthodes les plus utilisées (Fig. 2.1) pour déterminer les seuils de segmentation.

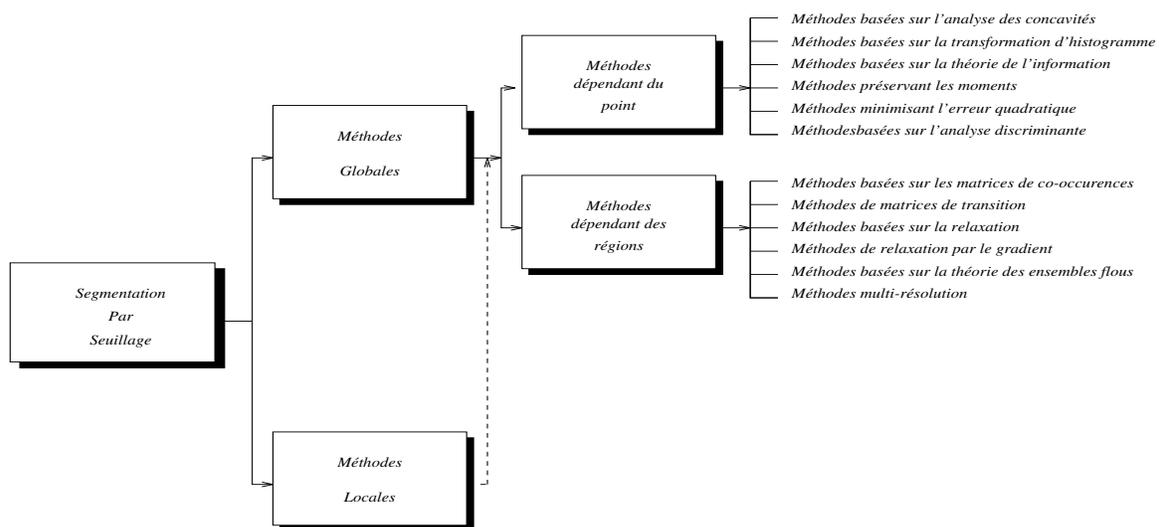


FIG. 2.1 - Classification des méthodes par seuillage

### 2.2.1.2.2 Approches globales

Les méthodes de segmentation par seuillage global sont des méthodes qui utilisent la même valeur de niveau de gris pour seuiller l'image entière. En général, une méthode de seuillage consiste à déterminer la valeur optimale du seuil  $s^*$  en se basant sur certains critères. Si  $s^*$  est déterminé uniquement à partir des niveaux de gris de chaque pixel, alors on parlera de méthodes de seuillage dépendant du point. Si  $s^*$  est déterminé à partir de propriétés locales sur un voisinage de chaque point, alors on parlera de méthodes de seuillage dépendant de la région (ici la région prend le sens de voisinage). Nous classerons les différentes méthodes suivant ce critère.

### □ Méthodes dépendant du point

Pour des images composées d'un faible nombre de régions de luminances différentes et bien contrastées, l'histogramme des niveaux de gris est multi-modal. La segmentation de l'image consiste alors à rechercher dans l'histogramme des luminosités la ou les vallées qui séparent ses modes. Les classes sont alors définies par les intervalles entre les vallées [272].

### • Méthodes basées sur l'analyse des concavités de l'histogramme

Lorsque les vallées sont mal définies à partir de l'histogramme des niveaux de gris, ce qui est souvent le cas, il est préférable de choisir le seuil en analysant les concavités de l'histogramme [301]. Afin de déterminer les concavités de l'histogramme  $h$ , on construit l'histogramme convexe  $\bar{h}$ . C'est le plus petit polynôme convexe de  $\bar{h}$  qui contient  $h$ . Les concavités de  $h$  sont déterminées à partir de la différence  $h - \bar{h}$ . Soit  $\bar{h}(g_i)$  la valeur de  $\bar{h}$  pour le niveau de gris  $g_i$ . Les choix de seuils possibles sont les niveaux de gris pour lesquels  $\bar{h}(g_i) - h(g_i)$  possède un maximum local. Cependant ces maxima ne correspondent pas tous à de bonnes valeurs pour le seuillage, car de grandes concavités peuvent être introduites par du bruit. Rosenfeld et De La Torre [301] appellent de telles concavités des "erreurs". Afin d'éliminer les maxima causés par de fausses concavités une mesure de compensation est introduite:

$$E_i = \left\{ \sum_{j=g_0}^{g_{i-1}} h(j) \right\} - \left\{ \sum_{j=g_i}^{g_{N_G-1}} h(j) \right\} \quad (2.2)$$

$E_i$  évalue la pondération de l'histogramme pour le niveau de gris  $g_i$ . Pour de fausses concavités, qui apparaissent d'un seul côté de l'histogramme, les valeurs de  $E_i$  sont faibles. Ainsi on peut ignorer les maxima de  $\bar{h} - h$  lorsque  $E_i$  est petite. Les maxima restant sont des seuils possibles pour le seuillage, mais ne sont pas optimaux, car les niveaux de gris proches de ces maxima pourraient aussi donner de bons résultats.

### • Méthodes basées sur la transformation d'histogramme

Des méthodes ont été proposées pour améliorer la dynamique de l'histogramme afin de mieux former ses vallées et ses modes. Une solution consiste à déterminer l'histogramme des niveaux de gris en ne prenant en compte que les points à faible gradient ou laplacien [350].

Dans [347], Watanabe suggère de sélectionner une valeur de seuil qui maximise la somme des gradients calculée sur tous les points dont le niveau de gris est égal à la valeur du seuil. Cette méthode a inspiré celle de Kohler [195] qui en plus est basée sur l'idée suivante: le seuil optimal pour la segmentation de l'image est celui qui détecte le plus grand nombre de contours à fort contraste et le plus petit nombre de contours à faible contraste. On cherche alors à maximiser l'influence des contours bien contrastés et à minimiser celle des contours peu contrastés en calculant le seuil  $s$  pour lequel la fonction suivante est maximale.

$$CM(s) = \frac{C(s)}{N(s)} \quad (2.3)$$

$N(s)$  dénote le nombre de points contours détectés par l'application du seuil  $s$ :

$$N(s) = \sum_{(i,j) \in A} p(g_i, g_j, s) \quad (2.4)$$

avec  $A$  est l'ensemble des couples de points adjacents  $(i, j)$  d'intensités respectives  $g_i$  et  $g_j$ , avec  $g_i \leq g_j$ , et:

$$p(g_i, g_j, s) = \begin{cases} 1 & \text{si } (g_i \leq s \leq g_j) \\ 0 & \text{sinon} \end{cases} \quad (2.5)$$

$C(s)$  correspond au contraste total détecté par l'application du seuil  $s$ :

$$C(s) = \sum_{(i,j) \in A} c(g_i, g_j, s) \quad (2.6)$$

$$c(g_i, g_j, s) = \begin{cases} \min(s - g_i, g_j - s) & \text{si } (g_i \leq s \leq g_j) \\ 0 & \text{sinon} \end{cases} \quad (2.7)$$

$CM(s)$  représente le contraste moyen observé aux abords des contours détectés par le seuil  $s$ .

La valeur optimale du seuil  $s^* = s$  est telle que  $CM(s^*) = \max_{s=1,2,\dots,N_G-1} CM(s)$ .

### • Méthodes basées sur la théorie d'information

D'autres méthodes de seuillage basées sur la théorie d'information ont été proposées. Certaines de ces méthodes utilisent uniquement l'entropie de l'histogramme [200] [287], d'autres tiennent compte de la distribution spatiale des niveaux de gris de l'image en introduisant des entropies d'ordres supérieurs [1] [260].

#### – La méthode de Pun

Nous allons décrire deux algorithmes basés sur la théorie d'information proposés par Pun [286] [287].

Soit  $s$  la valeur du seuil, deux mesures d'entropie a posteriori sont définies:

$$H'_b = - \sum_{i=0}^s p_i \ln p_i \quad H'_w = - \sum_{i=s+1}^{N_G-1} p_i \ln p_i \quad (2.8)$$

où  $p_i$  correspond à la probabilité d'occurrence du niveau de gris  $g_i$ , estimée par le rapport  $h(g_i)/N_G$ .  $H'_b$  et  $H'_w$  sont considérées comme les entropies associées aux pixels blanc et noir

après seuillage. Connaissant l'entropie de l'histogramme des niveaux de gris, Pun propose un algorithme qui détermine le seuil optimal en maximisant la bande limite de l'entropie a posteriori:

$$H' = H'_b + H'_w \quad (2.9)$$

Maximiser  $H'$  revient à maximiser la fonction d'évaluation  $F(s)$  telle que [254]:

$$F(s) = \frac{H_s}{H_T} \frac{\ln P_s}{\ln \max p_0, \dots, p_s} + \left[ 1 - \frac{H_s}{H_T} \right] \frac{\ln(1 - P_s)}{\ln \max p_{s+1}, \dots, p_{N_G-1}} \quad (2.10)$$

avec

$$H_s = - \sum_{i=0}^s p_i \ln p_i, \quad H_T = - \sum_{i=0}^{N_G-1} p_i \ln p_i \quad (2.11)$$

et

$$P_s = \sum_{i=0}^s p_i \quad (2.12)$$

Dans son second algorithme, Pun [286] propose d'utiliser le coefficient d'anisotropie  $\alpha$  donné par

$$\alpha = \frac{\sum_{i=0}^m p_i \ln p_i}{\sum_{i=0}^{N_G-1} p_i \ln p_i} \quad (2.13)$$

où  $m$  est le plus petit entier tel que

$$\sum_{i=0}^m p_i \geq 0.5 \quad (2.14)$$

le seuil optimal  $s^*$  est tel que

$$\sum_{i=0}^{s^*} p_i = \begin{cases} 1 - \alpha & \text{si } \alpha \leq 0.5 \\ \alpha & \text{si } \alpha > 0.5 \end{cases} \quad (2.15)$$

Cependant, Kapur et al. [200] ont montré que cet algorithme donnait pour toutes les images un seuil optimal  $s^* \geq m$ , donc introduisait inutilement un biais. Aucune cohérence spatiale n'est imposée aux objets détectés, ce qui dans le cas d'images bruitées crée des contours disjoints et une segmentation excessive des images.

### – La méthode de Kapur

Dans cette méthode [200], les deux distributions de probabilité (une pour les objets et une autre pour le fond) découlent de l'histogramme des niveaux de gris comme suit:

$$\frac{P_0}{P_s}, \frac{P_1}{P_s}, \dots, \frac{P_{s-1}}{P_s} \quad (2.16)$$

et

$$\frac{P_{s+1}}{1 - P_s}, \frac{P_{t+2}}{1 - P_s}, \dots, \frac{P_{N_G-1}}{1 - P_s} \quad (2.17)$$

où  $s$  est la valeur du seuil et  $P_s$  défini en 2.12.

On définit:

$$H_b(s) = - \sum_{i=0}^s \frac{P_i}{P_s} \ln \left( \frac{P_i}{P_s} \right) \quad (2.18)$$

$$H_w(s) = - \sum_{i=s+1}^{N_G-1} \frac{P_i}{1 - P_s} \ln \left( \frac{P_i}{1 - P_s} \right) \quad (2.19)$$

Alors, le seuil optimal  $s^*$  est défini comme le niveau de gris maximisant  $H_b(s) + H_w(s)$ :

$$s^* = \arg \max_{s \in G} \{H_b(s) + H_w(s)\} \quad (2.20)$$

### – La méthode de Johanssen et Bille

Cette méthode [177] utilise l'entropie de l'histogramme des niveaux de gris d'une image numérique. Ils séparent l'ensemble des niveaux de gris en deux parties de manière à minimiser l'interdépendance entre elles.

La méthode conduit au calcul de la valeur du seuil  $s^*$  à partir de la relation:

$$s^* = \arg \max_{s \in G} S(s) + \bar{S}(s) \quad (2.21)$$

avec

$$S(s) = \ln(P_s) - \frac{1}{P_s} \left[ p_s \ln p_s + \left( \sum_{i=0}^{s-1} p_i \right) \ln(P_s) \right] \quad (2.22)$$

et

$$\bar{S}(s) = \ln(1 - P_s) - \frac{1}{(1 - P_s)} \left[ p_s \ln p_s + \left( \sum_{i=s+1}^{N_G-1} p_i \right) \ln \left( \sum_{i=s+1}^{N_G-1} p_i \right) \right] \quad (2.23)$$

Ces méthodes d'entropie utilisent l'intégrale de l'histogramme comme fonction de distribution. Les frontières entre régions correspondent au passage par zéro de l'intégrale. La recherche des contours se fait en maximisant l'entropie très sensible aux défauts d'homogénéité d'éclairage, ce qui crée des frontières très chahutées et une segmentation excessive des images.

### • Méthode préservant les moments

Dans cette méthode [333], les valeurs des seuils sont calculées en tenant compte du fait que les trois moments de l'image sont préservés entre l'image initiale et l'image binarisée.

Le  $i$ ème moment  $\mu_i$  est calculé comme suit

$$\mu_i = \frac{1}{N_G} \sum_{g=0}^{N_G-1} g^i h(g), \quad i = 1, 2, 3, \dots \quad (2.24)$$

$N_G$  est le nombre total de pixels dans l'image.

La valeur du seuil  $s^*$  est choisi comme le niveau de gris représentant  $p_0$ , qui est donné par

$$p_0 = \frac{z - \mu_1}{(c_1^2 - 4c_0)^{1/2}} \quad (2.25)$$

avec

$$c_0 = \frac{\mu_1\mu_3 - \mu_2^2}{\mu_2 - \mu_1^2} \quad c_1 = \frac{\mu_1\mu_2 - \mu_3}{\mu_2 - \mu_1^2} \quad (2.26)$$

et

$$z = \frac{1}{2}\{(c_1^2 - 4c_0)^{1/2} - c_1\} \quad (2.27)$$

### • Méthode basée sur l'analyse discriminante

Otsu propose dans [254] une méthode basée sur l'analyse discriminante. Cette méthode tente de segmenter l'image en  $m$  classes en maximisant un critère de séparabilité entre classes. Dans le cas de deux classes  $\Omega_0$  et  $\Omega_1$ , les relations suivantes permettent de déterminer le seuil  $s$  de séparation entre  $\Omega_0$  et  $\Omega_1$  (objet et fond).

Soit  $p_i$  la probabilité d'occurrence de niveau de gris  $g_i$ ,  $i = 0, 1, \dots, N_G - 1$ .

Soient  $\mu$  et  $\sigma$  respectivement la valeur moyenne et l'écart type des intensités de l'image.

Les probabilités pour qu'un point appartienne à  $\Omega_0$  ou  $\Omega_1$  sont données respectivement par

$$P_s \quad \text{et} \quad 1 - P_s \quad (2.28)$$

Les valeurs moyennes des intensités calculées dans  $\Omega_0$  et  $\Omega_1$  sont données par

$$\mu_T = \sum_{i=0}^s ip_i \quad \text{et} \quad 1 - \mu_T \quad (2.29)$$

Après normalisation de ces valeurs par la probabilité d'appartenance à chaque classe on obtient

$$\mu_0 = \frac{\mu_T}{P_s} \quad \text{et} \quad \mu_1 = \frac{1 - \mu_T}{1 - P_s} \quad (2.30)$$

On en déduit alors la variance qui sépare  $\Omega_0$  de  $\Omega_1$ :

$$\sigma_B^2(s^*) = P_s(\mu_0 - \mu)^2 + (1 - P_s)(\mu_1 - \mu)^2 \quad (2.31)$$

Le seuil  $s^*$  qui maximise la variance interclasses est alors fourni par la relation

$$\sigma_B^2(s^*) = \max_{s=1, \dots, N_G-1} \{\sigma_B^2(s)\} \quad (2.32)$$

Cette méthode permet d'étiqueter les pixels d'une image en deux classes: objets et fond. Le seuil ainsi calculé assure une satisfaisante localisation des frontières des régions, pour

une grande variété d'images.

Dans [299], Reddi et al. proposent une généralisation de cette méthode pour la détection de plusieurs seuils. D'autres méthodes s'appuyant sur l'analyse discriminante ont été proposées. Parmi ces méthodes on peut citer celle des K-means [113] et celle des centres mobiles [323].

### • Méthode minimisant l'erreur quadratique

Dans [191], l'histogramme des niveaux de gris est considéré comme une estimée de la fonction de densité de probabilité  $p(g)$  d'un mélange de population comprenant les niveaux de gris des objets et du fond. On peut supposer que chacune des deux composantes  $p(g|\Omega_i)$  du mélange est une distribution normale avec une moyenne  $\mu_i$ , un écart type  $\sigma_i$  et d'une probabilité a priori  $P_i$ :

$$p(g) = \sum_{i=1}^2 P_i p(g|\Omega_i) \quad (2.33)$$

avec

$$p(g|\Omega_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(g-\mu_i)^2}{2\sigma_i^2}\right) \quad (2.34)$$

La valeur du seuil optimal peut être déterminée en résolvant l'équation quadratique suivante:

$$\frac{(g-\mu_1)^2}{\sigma_1^2} + \ln \sigma_1^2 - 2 \ln P_1 = \frac{(g-\mu_2)^2}{\sigma_2^2} + \ln \sigma_2^2 - 2 \ln P_2 \quad (2.35)$$

Souvent, les paramètres  $\mu_i$ ,  $\sigma_i^2$  et  $P_i$  ( $i = 1, 2$ ) de la densité de probabilité  $p(g)$  associée à l'image ne sont pas connus. Pour surmonter la difficulté de l'estimation des paramètres inconnus, Kittler et Illingworth [191] ont introduit la fonction  $J(s)$  donnée par

$$J(s) = 1 + 2P_1(s) \ln \sigma_1(s) + P_2(s) \ln \sigma_2(s) - 2P_1(s) \ln P_1(s) + P_2(s) \ln P_2(s) \quad (2.36)$$

avec

$$P_1 = \sum_{g=0}^s h(g), \quad P_2 = \sum_{g=s+1}^{N_G-1} h(g) \quad (2.37)$$

et

$$\mu_1 = \frac{\sum_{g=0}^s h(g)g}{P_1(s)}, \quad \mu_2 = \frac{\sum_{g=s+1}^{N_G-1} h(g)g}{P_2(s)} \quad (2.38)$$

et

$$\sigma_1^2(s) = \frac{\sum_{g=0}^s (g-\mu_1(s))^2 h(g)}{P_1(s)}, \quad \sigma_2^2(s) = \frac{\sum_{g=s+1}^{N_G-1} (g-\mu_2(s))^2 h(g)}{P_2(s)} \quad (2.39)$$

en minimisant  $J(s)$ , le seuil optimal est défini par

$$s^* = \arg \min_{s \in G} J(s) \quad (2.40)$$

Le seuil déterminé par cette méthode ne permet pas une détection précise des régions, lorsque les distributions de niveaux de gris des deux classes sont légèrement mélangées.

Un des inconvénients des méthodes de seuillage dépendant du point est qu'elles font intervenir essentiellement les statistiques du premier ordre des niveaux de gris de l'image, c'est à dire l'histogramme des niveaux de gris. On n'obtient alors aucune information spatiale de la répartition des niveaux de gris. La plupart des méthodes décrites ci-après sont basées sur les statistiques du second ordre des niveaux de gris.

### □ Méthodes dépendant des régions

#### • Méthodes basées sur les matrices de co-occurrence

Les matrices de co-occurrence ont été introduites par Haralick [168] pour l'analyse des textures (une étude détaillée est présentée au paragraphe 2.2.3.2.2). En général, une matrice de co-occurrence  $M_{(d,\theta)}$  est une estimée de la densité de probabilité conjointe de rencontrer deux niveaux de gris  $i$  et  $j$  séparés par une distance  $d$  et avec une orientation  $\theta$ . Le nombre de ces matrices n'est pas limité, puisqu'elles dépendent du choix de  $d$  et  $\theta$ . Dans [11], Ahuya et Rosenfeld ont défini la matrice de co-occurrence des niveaux de gris suivante:

$$M = M_{(1,0)} + M_{(1,\pi/2)} + M_{(1,\pi)} + M_{(1,3\pi/2)} \quad (2.41)$$

Ainsi l'élément  $(i, j)$  de  $M$  est la fréquence d'apparition du niveau de gris  $i$  parmi les 4 voisins de niveau de gris  $j$ .

Du fait d'une certaine homogénéité des pixels appartenant aux objets ou au fond, ces pixels devraient contribuer principalement aux éléments de la diagonale de  $M$ . Aussi les pixels proches des contours devraient contribuer aux éléments de la sous-diagonale et de la sur-diagonale de  $M$  car les niveaux de gris changent au voisinage d'un contour. Ainsi la matrice  $M$  peut être utilisée pour calculer deux nouveaux histogrammes:

- un histogramme construit à partir des éléments de la diagonale de  $M$ . Cet histogramme devrait avoir une vallée profonde entre les niveaux de gris des objets et ceux du fond.
- un histogramme construit à partir des éléments hors-diagonale de  $M$ . Cet histogramme devrait posséder un mode étroit entre les niveaux de gris des objets et ceux du fond.

Le seuil de l'image est le niveau de gris qui correspond le mieux à la vallée du premier histogramme et donc au pic du second histogramme, si on les superpose.

Dans [277], Pal et al. définissent une mesure de contraste et une mesure d'homogénéité à partir des matrices de co-occurrence. La mesure de contraste est fondée sur la réponse logarithmique du système visuel humain. Une troisième mesure, dite mesure composite, est définie à partir des deux précédentes. Le niveau de gris  $s^*$  qui maximise cette mesure composite est considéré comme le contour (ou le seuil) entre deux régions. Cette méthode

permet d'obtenir une segmentation au sens des niveaux de gris, et de supprimer les seuils indésirables qui auraient été obtenus par une division classique par histogrammes locaux.

Nous pouvons également citer la méthode de Houzelle et Giraudon [160] qui utilisent un algorithme de croissance de régions pour segmenter des images au sens des niveaux de gris. Le critère d'homogénéité utilisé dans la croissance repose sur l'étude des caractéristiques extraites de la diagonale de la matrice de co-occurrence caractérisant les régions homogènes au sens des niveaux de gris présentes dans la fenêtre étudiée.

### • Méthode des matrices de transition

Les matrices de transition sont similaires aux matrices de co-occurrence. Elles sont utilisées pour définir deux mesures d'interaction pour le calcul du seuil. Le seuil optimal est déterminé en minimisant ces deux mesures. En prenant les notations du paragraphe précédent, les matrices de transition définies dans [117] sont:

$$T_h = M_{(1,0)}, \quad T_v = M_{(1,3\pi/2)}, \quad T_{vh} = T_v + T_h \quad (2.42)$$

N'importe quelle matrice de transition définie ci-dessus peut être utilisée pour la sélection d'un seuil. Il est mentionné dans [117] que le résultat final est très peu sensible au choix des matrices de transition.

Soient  $T_{ij}$  l'élément  $(i, j)$  de  $T$  et  $s$  le seuil qui sépare l'ensemble des niveaux de gris en deux classes  $\Omega_0 = \{0, 1, \dots, s\}$  et  $\Omega_1 = \{s + 1, s + 2, \dots, N_G - 1\}$ , alors  $T$  peut être décomposée en 4 régions définies par les paramètres suivants

$$a = \sum_{i=0}^s \sum_{j=0}^s T_{ij}, \quad b = \sum_{i=s+1}^{N_G-1} \sum_{j=s+1}^{N_G-1} T_{ij}, \quad c = \sum_{i=0}^s \sum_{j=s+1}^{N_G-1} T_{ij}, \quad d = \sum_{i=s+1}^{N_G-1} \sum_{j=0}^s T_{ij} \quad (2.43)$$

alors  $a, b, c$  et  $d$  représentent le nombre total de transitions correspondant respectivement à l'intérieur de  $\Omega_0$ , à l'intérieur de  $\Omega_1$ , entre  $\Omega_0$  et  $\Omega_1$ , et entre  $\Omega_1$  et  $\Omega_0$ . Les probabilités jointes et conditionnelles des transitions entre  $\Omega_1$  et  $\Omega_0$  peuvent être estimées par

$$P_j(s) = \frac{c + d}{a + b + c + d} \quad (2.44)$$

$$P_c(s) = 1/2 \left\{ \frac{c}{a + c} + \frac{d}{b + d} \right\} \quad (2.45)$$

Deravi et Pal [117] appellent  $P_j(s)$  et  $P_c(s)$  des mesures d'interaction. Ils montrent que  $P_j(s)$  est similaire à la mesure décrite dans [361] et que  $P_c(s)$  ne dépend pas directement de l'histogramme. Le seuil optimal  $s^*$  est obtenu en maximisant l'une ou l'autre des mesures d'interaction.

Chanda [57] utilise une mesure de contraste moyen à partir d'une matrice de co-occurrence non symétrique définie pour la translation  $(1, 3\pi/2)$ .

La mesure de contraste moyen pour le niveau de gris  $s$  est alors

$$C(s) = \frac{\sum_{i=0}^s \sum_{j=s+1}^{N_G-1} (i-j)^2 T_{ij}}{\sum_{i=0}^s \sum_{j=s+1}^{N_G-1} T_{ij}} + \frac{\sum_{j=0}^s \sum_{i=s+1}^{N_G-1} (i-j)^2 T_{ij}}{\sum_{j=0}^s \sum_{i=s+1}^{N_G-1} T_{ij}} \quad (2.46)$$

ou encore, en reprenant les notations précédentes:

$$C(s) = \sum_{i=0}^s \sum_{j=s+1}^{N_G-1} \frac{(i-j)^2 T_{ij}}{c} + \sum_{j=0}^s \sum_{i=s+1}^{N_G-1} \frac{(i-j)^2 T_{ij}}{d} \quad (2.47)$$

Cette fonction est analogue à un calcul de variance. Considérons le premier terme,  $T_{ij}$  compte le nombre de fois où le niveau de gris  $i$  a succédé au niveau de gris  $j$  sur la verticale. Ainsi,  $T_{ij}/c$  est proportionnelle à une fréquence de transition entre  $i$  et  $j$ . Quant au facteur  $(i-j)^2$ , comme dans le calcul d'une variance, il correspond à une dispersion, c'est-à-dire à l'amplitude de cette transition. Alors, le premier terme (resp. second terme) de  $C(s)$  est analogue à l'amplitude moyenne de transition sur  $c$  (resp. sur  $d$ ). Puisque  $c$  et  $d$  représentent le nombre total de transitions inter-région, les maxima de  $C(s)$  donneront les seuils  $s$ .

#### • Méthodes basées sur la relaxation

L'idée de la relaxation a été introduite par Southwell [317] pour améliorer la convergence de la solution récursive des systèmes d'équations linéaires.

Dans le cas du seuillage d'images, la relaxation est appliquée de la façon suivante: les pixels d'une image sont d'abord séparés en deux classes; "clair" et "foncé", suivant leurs niveaux de gris. Ainsi la probabilité de chaque pixel est ajustée suivant la probabilité de ses voisins. Ce procédé d'ajustement est réitéré tant que la probabilité d'être clair (resp. foncé) reste importante pour des pixels appartenant aux régions claires (resp. foncées). Ces méthodes sont basées sur deux étapes [80] [124] [257] [300] [371]: la classification initiale et le réajustement itératif des probabilités.

Les méthodes de relaxation sont très coûteuses en temps de calcul et en place mémoire. Elles sont sensibles aux zones d'ombre présentes dans un fond continu.

#### • Méthodes basées sur la notion des ensembles flous

Les méthodes floues de classification ont été étudiées par plusieurs auteurs [51] [77] [91] [116] [212] [239] [282] [366]. D'une façon générale, la classification floue permet de moduler la décision finale d'étiquetage des points en fonction des connaissances a priori sur le contenu de l'image. Les points ayant un degré d'appartenance supérieur à un seuil sont qualifiés de sûrs, les autres sont classés ambigus.

Dans le cas du seuillage, on calcule les valeurs d'appartenance de chaque élément à chacune des classes itérativement jusqu'à l'optimisation d'un critère. De cette manière chaque élément peut appartenir à plusieurs classes avec un degré plus ou moins fort, permettant de conserver longtemps un volume d'information important sans avoir à prendre la décision prématurément.

Parmi les algorithmes de classification floue on peut citer le "Fuzzy-C Means" [29]. Cet algorithme est basé sur un critère quadratique qui sert à classifier des formes en retardant la décision jusqu'à obtention d'une quantité d'information suffisante. Le critère quadratique à minimiser est la somme pondérée, pour toutes les classes formées, des écarts quadratiques intra-classes.

Soit  $X = \{x_1, x_2, \dots, x_M\}$  un ensemble fini de formes de dimension  $M$ ,  $c$  un entier appartenant à  $\{2, 3, \dots, M\}$  représentant le nombre de classes et  $U = \mu_{ik}$  une partition floue de  $X$  en  $c$  classes, où chaque  $\mu_{ik}$  est le degré d'appartenance  $\mu_k(x_i)$  de la forme  $x_i$  à la classe  $k$ . Supposons  $\sum_{k=1}^c \mu_{ik} = 1 \forall i$ , alors le critère quadratique de classification  $J_m$  est défini par:

$$J_m(U, v) = \sum_{i=1}^M \sum_{k=1}^c (\mu_{ik})^m (d_{ik})^2 \quad (2.48)$$

où  $m$  est le facteur de flou ( $1 \leq m < \infty$ ),  $d_{ik}$  est une distance quelconque entre la forme  $x_i$  et le vecteur  $v = (v_1, v_2, \dots, v_c)$  des centres des classes. Lorsque  $m$  tend vers 1 et  $d$  est la distance euclidienne, l'algorithme "Fuzzy-C Means" fournit la partition binaire optimale.

Il n'y a pas de règles pour fixer la valeur de  $m$  car il n'existe pas de base théorique pour son optimisation. Cela permet de mettre en valeur l'ambiguïté existante dans l'ensemble à classer, ou au contraire, de l'atténuer. Le facteur de flou  $m$  interfère sur deux caractéristiques de l'algorithme: la rapidité de convergence décroît avec l'augmentation de  $m$ , en même temps que l'apport de chaque élément dans le calcul des centres des classes décroît. Pour les valeurs de  $m$  supérieures à deux, les partitions tendent vers le centre de gravité de l'espace des partitions floues, c.à.d. vers la partition  $U = [1/c]$ , ce qui n'offre pas de réel intérêt. Pour les valeurs de  $m$  proches de 1, les degrés d'appartenance sont très stricts (proches des valeurs binaires) et ne traduisent pas l'ambiguïté d'affectation des formes de  $X$ . Enfin, les valeurs de  $m$  prises dans l'intervalle  $[1.5, \dots, 2]$  permettent d'obtenir des résultats intéressants et d'interprétation aisée.

Le nombre d'itérations de l'algorithme "Fuzzy-C Means" augmente avec la précision demandée sur les valeurs des centres (ou des degrés d'appartenance, si le test d'arrêt est fait en fonction de celles-ci). Pour les formes entières (les niveaux de gris, par exemple) un seuil de convergence de 0.01 donne des erreurs inférieures à 0.5%, ce qui apporte un bon compromis entre la rapidité de convergence et la précision des résultats. En classification floue, l'appartenance d'un élément à une classe est d'autant plus forte que son degré d'appartenance à cette classe est proche de 1 et que ses degrés d'appartenance aux autres classes sont proches

de 0. La meilleure classification floue réalisable est donc la classification qui correspond le plus à une partition binaire.

#### □ Autres méthodes

De nombreuses autres méthodes peuvent être trouvées dans la littérature, Par exemple, Panda suggère une méthode pour laquelle le seuil appliqué à un pixel dépend à la fois du niveau de gris et du gradient en ce point. Des résultats et des comparaisons sont donnés dans [256].

Dans [238], Morrin utilise le gradient des niveaux de gris afin de changer l'échelle des niveaux de gris de l'image, pour avoir une meilleure résolution et un meilleur contraste pour le seuillage.

#### □ Conclusion sur la segmentation par seuillage global

Malgré les améliorations qui peuvent être apportées aux données de l'image, les techniques de segmentation par seuillage global s'avèrent non satisfaisantes pour une grande variété d'images. Une des raisons de ce constat est liée au manque d'utilisation de relations spatiales le long du processus de segmentation et à une superposition d'informations due à l'importance de la taille des données étudiées.

Pour pallier en partie cette lacune, d'autres approches dites de seuillage local ont été développées, c'est le propos du paragraphe suivant.

### 2.2.1.2.3 Approches locales

Dans cette famille de méthodes on cherche à déterminer en chaque point un seuil dépendant de l'histogramme de répartition des luminances de son voisinage.

Une méthode de segmentation par seuillage local consiste à subdiviser l'image en blocs de tailles égales (généralement de taille 32x32). Puis un seuil optimal est calculé pour chacun des blocs en utilisant une des méthodes citées précédemment [36]. Une technique classique pour déterminer les seuils localement est présentée dans [81].

Une autre méthode de seuillage local a été suggérée par Fernando et Monro [127] pour segmenter des images d'angiogramme prises à partir de rayons X (les méthodes de seuillage global n'ont pas donné de résultats satisfaisants pour ces images, car elles sont essentiellement unimodales avec un pic très étroit). Pour cette méthode, l'image est divisée en 16 sous images, qui ne se chevauchent pas, et la technique de seuillage de l'entropie de Pun [286] est employée pour déterminer la valeur du seuil pour chaque sous image. Finalement, l'image entière est seuillée après un filtrage passe-bas qui élimine les discontinuités dues aux

frontières des sous images voisines (effets de bords).

Dans [75], Chehikian propose deux algorithmes pour résoudre le problème cité. Le premier est fondé sur l'analyse multi-fréquence de l'image pour déduire l'appartenance ou non du pixel à l'objet, le second procède par l'apprentissage d'un seuil local et la classification des pixels est basée sur ce seuil.

Dans *l'approche multi-résolution*, une séquence de convolution par un filtre gaussien permet, de manière contrôlée, une réduction de la fréquence maximale contenue dans l'image. Lorsque cette fréquence maximale a été réduite d'un facteur de deux, on procède à un sous échantillonnage du même facteur. Une séquence de filtrage permet d'obtenir à partir d'une image, une suite d'images de résolutions différentes. L'ensemble des images ainsi obtenues constitue une "pyramide multi-résolution" permettant d'avoir d'une même scène un ensemble de "vues", depuis la plus globale (image au sommet de la pyramide), jusqu'à la plus détaillée (image à la base de la pyramide). En procédant à une remise à l'échelle de chacune des images par une expansion-interpolation, la pyramide multi-résolution est transformée en une représentation multi-fréquence. A partir de cette représentation multi-fréquence, chaque pixel  $(i, j)$  est étiqueté "blanc" ou "noir" par comparaison aux différents laplaciens précédemment calculés.

L'exploitation des propriétés du laplacien de haute fréquence, permet d'étiqueter les pixels intérieurs et extérieurs proches du contour des objets, les laplaciens de plus basses fréquences étant utilisés pour remplir l'intérieur des objets par couches successives. Ils servent à établir un lien entre les pixels intérieurs situés au voisinage d'un contour et les pixels intérieurs éloignés du contour. Il est donc possible d'étiqueter les pixels par comparaison à un seuil à condition de connaître les niveaux de gris caractéristiques de l'intérieur et de l'extérieur des objets. Le laplacien de haute fréquence permet de connaître ces niveaux de gris.

L'avantage essentiel de cet algorithme réside dans le fait qu'à aucun moment un seuil quelconque n'est utilisé (il faut cependant décider, en présence du bruit dû en particulier à la numérisation, du signe du laplacien). Un inconvénient lié à cet algorithme résulte de sa capacité à binariser des images à très faible contraste. A la limite, il binarise de la même façon des objets réels et des objets fictifs dus à des variations du niveau de gris du fond. Il est toujours possible de contrôler la binarisation en intervenant sur le codage de laplaciens négatifs au moyen d'un seuil  $s$ . Il est néanmoins très difficile de lier une différence de niveau de gris et la valeur d'une dérivée seconde faisant intervenir à la fois une différence de niveau de gris et une étendue sur laquelle cette différence est évaluée. Ce sont ces raisons qui ont amené Chéhikian à proposer un second algorithme qui procède par apprentissage local d'un seuil. Cette méthode s'est avérée très efficace sur une classe d'images présentant des objets sombres sur un fond clair ou vice-versa. Cet algorithme nécessite néanmoins le choix d'un contraste minimal, qu'il est possible de définir suivant l'application.

### 2.2.1.2.4 Évaluation des techniques de seuillage

Dans leur papier [319], Sahoo et al. se sont inspirés des travaux de Nazif et Levine [215] pour réaliser une étude comparative des méthodes de seuillage d’histogramme suivantes: la méthode basée sur l’analyse discriminante [254], les trois méthodes d’entropie [177], [200], [286], les méthodes utilisant les matrices de co-occurrence [11], la méthode d’analyse des concavités de l’histogramme [301], la méthode utilisant les matrices de transition [117], la méthode préservant les moments [333] et la méthode basée sur l’erreur quadratique minimale [191]. L’évaluation s’est faite à partir de deux critères; un, portant sur l’uniformité des régions et un autre, portant sur la détection des contours.

La mesure d’uniformité utilisée ici est représentée, pour une valeur de seuil  $s$ , par la relation suivante:

$$U(s) = 1 - \frac{\sigma_1^2 + \sigma_2^2}{C} \quad (2.49)$$

avec

$$\sigma_i^2 = \sum_{(x,y) \in R_i} (I(x,y) - \mu_i)^2 \quad (2.50)$$

où

$R_i$  correspond à la  $i$ ème région segmentée,

$I(x, y)$  est le niveau de gris du pixel  $(x, y)$ ;

$$\mu_i = \frac{\sum_{(x,y) \in R_i} I(x,y)}{A_i},$$

$A_i$  est le nombre de pixels dans la région  $i$ ,  $i = 1, 2, \dots$

$C$  est un facteur de normalisation.

Levine et Nazif [215] ont introduit un facteur de pondération dans le calcul de la mesure d’uniformité  $U$  pour des images qui possèdent plus de deux régions.

La mesure de contours  $M_c$  est utilisée pour détecter les contours des objets dans une image. Cette mesure est calculée de la façon suivante :

A chaque pixel est associée la valeur du gradient généralisé  $\Delta(x, y)$ .

Si le pixel  $(x, y)$  a un niveau de gris supérieur à la moyenne des niveaux de gris de ses voisins alors on affectera le signe “+” au gradient généralisé  $\Delta(x, y)$ , sinon on lui affectera le signe “-”.

La mesure de  $M_c$  est alors donnée par la formule suivante:

$$M_c = \frac{\sum Sgn(I(x, y) - I_{N(x,y)}^-) \Delta(x, y) Sgn(I(x, y) - s)}{C} \quad (2.51)$$

où  $I_{N(x,y)}^-$  représente la moyenne des valeurs du niveau de gris sur le voisinage  $N(x, y)$ .  $s$  est la valeur du seuil.  $C$  est le facteur de normalisation.

$$Sgn(x) = \begin{cases} 1 & \text{si } (x \geq 0) \\ -1 & \text{si } (x < 0) \end{cases} \quad (2.52)$$

Pour calculer le gradient généralisé  $\Delta(x, y)$  du pixel  $(x, y)$ , on applique la formule:

$$\Delta(x, y) = \left[ \sum_{k=1}^4 D_k^2 + \sqrt{2}D_1(D_3 + D_4) - \sqrt{2}D_2(D_3 - D_4) \right] \quad (2.53)$$

avec

$$D_1 = I(x + 1, y) - I(x - 1, y)$$

$$D_2 = I(x, y - 1) - I(x, y + 1)$$

$$D_3 = I(x + 1, y + 1) - I(x - 1, y - 1)$$

$$D_4 = I(x + 1, y - 1) - I(x - 1, y + 1)$$

Les différentes méthodes de seuillage ont été testées sur une image d'un "cameraman" qui présente un histogramme de niveaux de gris global bimodal et sur une image d'un "immeuble" qui n'a pas un histogramme de niveaux de gris franchement bimodal tout comme une troisième image d'un "portrait féminin". Les résultats de comparaison basés sur l'ensemble de ces critères ont montré que la méthode basée sur l'analyse discriminante donne les meilleurs résultats, pour les différentes images tests. La méthode de préservation des moments ainsi que la seconde méthode d'entropie (méthode de Kapur) donnent également de bons résultats. La méthode des matrices de co-occurrence préserve bien les contours et la méthode basée sur l'analyse discriminante sépare bien les régions uniformes.

Si on considère d'autres critères d'évaluation comme l'appréciation des images, les détails, les contours, les distorsions..., il s'avère que des méthodes peu performantes vis à vis des deux critères (uniformité  $U(s)$  et formes  $M_c$ ) donnent ici de bons résultats. Alors, à partir d'une évaluation visuelle de l'image binaire, tout en s'appuyant sur une évaluation objective à partir des deux critères d'uniformité et de formes, il est montré que la méthode préservant les moments, la méthode de Kapur et la méthode de Johannsen et Bille basées sur l'entropie, ainsi que la méthode d'Otsu basée sur l'analyse discriminante, sont de bonnes méthodes de seuillage. La méthode d'entropie de Pun, la méthode basée sur les concavités d'histogramme, et celle de l'erreur quadratique minimum, sont relativement moins bonnes pour les images tests qui ont été utilisées. Elles sont reconnues néanmoins pour donner de bons résultats, pour les applications pour lesquelles elles ont été développées.

### 2.2.1.2.5 Conclusions sur les techniques de seuillage

Nous avons passé en revue plusieurs méthodes de seuillage globales et locales. Pour la plupart des méthodes de seuillage local, l'image est subdivisée en petites sous-images et ensuite des techniques globales de seuillage sont fréquemment utilisées.

L'avantage principal des ces méthodes réside dans le fait que deux pixels quelconques de même niveau de gris sont traités d'une manière identique. Ces deux pixels auront donc le même degré d'appartenance à chacune des classes. Il n'est pas alors nécessaire de classer tous les pixels d'une image, mais simplement de classer les différentes valeurs de niveaux de gris qu'elle représente. La difficulté de seuillage des images est liée aux problèmes de l'in-homogénéité de la réponse du capteur, aux effets d'ombre, et à la nature des images à traiter.

Ces méthodes de seuillage donnent toutes des résultats satisfaisants suivant la classe d'images pour laquelle elles ont été développées. Cependant, elles ne sont pas toutes automatiques; certaines sont interactives et nécessitent un opérateur extérieur pour l'obtention d'un seuil optimal.

L'automatisation de la technique de seuillage présente encore des difficultés car un mauvais seuillage entraîne systématiquement une extraction erronée des paramètres caractérisant chaque région à modéliser (suppression de certains détails discriminants de l'objet pour son identification, mauvais tracé du contour,...).

Dans la section 3.3.1.1.4, une comparaison de la plupart des méthodes présentées est effectuée sur plusieurs types d'images en utilisant des critères objectifs basés sur l'uniformité des régions et la détection des contours.

### 2.2.1.3 Méthodes de segmentation par croissance de régions

#### 2.2.1.3.1 Introduction

Les inconvénients majeurs des méthodes basées sur le seuillage d'histogramme des luminances proviennent du fait qu'elles ne tiennent pas compte des relations spatiales des pixels. La segmentation par croissance de régions utilise aussi bien les relations spatiales entre les points que leurs propriétés. L'objectif est d'obtenir des régions significatives et homogènes au sens d'un critère d'homogénéité. Le problème fondamental associé à cette catégorie d'approches est justement lié à la définition du critère d'homogénéité utilisé dans le processus de partitionnement de l'image.

Avant d'aborder l'aspect technique de ce type de segmentation, nous rappelons la définition classique du processus en s'appuyant sur la définition proposée par Zucker [376].

La segmentation en régions d'une image  $I$  revient à réaliser sur  $I$  une partition en sous ensembles  $R_i (i = 1, \dots, N)$  tels que si  $\mathcal{P}$  est un prédicat sur un certain nombre de points, on ait :

1.  $I = \bigcup_{i=1}^N R_i$ ,
2.  $R_i \cap R_j = \emptyset \quad \forall (i, j) \text{ tel que } i \neq j$ ,
3. les sous ensembles  $R_i (i = 1, \dots, N)$  sont connexes,
4.  $\mathcal{P}(R_i) = \text{vrai}, \quad \forall i = 1 \dots N$ ,
5.  $\mathcal{P}(R_i \cap R_j) = \text{faux} \quad \forall (i, j) / i \neq j, R_i \text{ et } R_j \text{ contiguës.}$

On note qu'une segmentation qui vérifie les propriétés précédentes ne garanti pas dans tous les cas une partition unique de l'image. Pour une même image, le nombre de régions  $N$  est variable en fonction du prédicat  $\mathcal{P}$  et/ou des seuils de tolérance fixés en vue de leur formation. Les résultats d'une segmentation dépendent de l'ordre dans lequel les régions ont été parcourues. Généralement les algorithmes de segmentation ne donnent pas les mêmes résultats sur une image et sa transposée.

Le prédicat d'uniformité prend la forme suivante :

$\mathcal{P}(R) = \text{vrai} \quad \text{si un critère } C \text{ défini sur les attributs de l'ensemble des points de la région } R \text{ de l'image est vérifié.}$

Le processus général d'extraction de régions est généralement itératif. Il est initialisé par une partition de l'image en régions de base, puis de manière séquentielle ou récursive, ces régions sont traitées.

On peut globalement classer ces méthodes en deux catégories: celles qui font des agrégations de points et celles qui opèrent par fusion et/ou division d'ensembles de points (régions).

Le premier type de méthodes procède par comparaison des propriétés d'un point avec un point voisin et décide du regroupement des deux points si leurs propriétés sont suffisamment proches. Certaines méthodes d'agrégation de points fonctionnent selon le principe suivant: on utilise un graphe où chaque noeud correspond à un point. Deux noeuds sont reliés par un arc s'ils sont similaires. La segmentation correspond à l'extraction des composantes connexes du graphe. L'heuristique de similarité la plus simple postule la fusion de deux noeuds si leur différence de niveau de gris est négligeable. Un des reproches faits à ce type de méthodes est dû à l'utilisation de critères trop locaux et au nombre restreint d'informations que véhiculent les points à regrouper.

Des solutions ont été proposées pour remédier à ce problème en introduisant d'autres critères de regroupement relativement sophistiqués et opérant sur des ensembles de points: ce sont les techniques de la seconde catégorie de méthodes. La segmentation est effectuée en

analysant des régions ou des couples de régions adjacentes de tailles et de formes variées et la formation de régions est faite d'une manière séquentielle ou itérative si les conditions de division et/ou de fusion sont vérifiées.

Dans cette catégorie trois grandes familles de méthodes sont distinguées:

- les méthodes de division qui procèdent par division des régions de base en régions plus petites et de plus en plus homogènes. La division s'arrête quand les régions produites vérifient un certain critère d'homogénéité.
- les méthodes de fusion qui procèdent par croissance de régions de base. La fusion s'arrête lorsqu'une partition optimale au sens du critère d'homogénéité est atteinte.
- les méthodes de division-fusion qui combinent les deux méthodes précédentes.

Nous présentons, dans la suite de cette section, l'ensemble de ces différentes méthodes et illustrons chacune d'elles par des exemples.

### 2.2.1.3.2 Approches par fusion

L'idée consiste à exploiter une partition initiale de l'image constituée de petites régions. Puis ces régions sont fusionnées successivement jusqu'à ce que le critère de fusion ne soit plus vérifié.

Plusieurs règles de regroupement ont été proposées. Certaines de ces règles mettent en jeu:

- des propriétés statistiques telles que la moyenne ou la variance des niveaux de gris des régions, le gradient moyen des frontières de régions, le contraste maximum des régions, ou d'autres statistiques locales qui expriment l'état de surface des régions, etc.
- des propriétés géométriques ou morphologiques telle que l'élongation ou la compacité des régions. Deux régions sont regroupées si par exemple un facteur de forme est conservé ou amélioré après leur fusion.

A titre d'exemple de ces méthodes, nous pouvons citer celle de Pong et al. [255] qui décrivent un processus de fusion qui exploite un modèle à facette développé par Haralick et Watson [172]. La segmentation est obtenue par regroupement de facettes ayant des propriétés similaires.

Dans [135] les auteurs suggèrent d'utiliser une succession de prédicats. Lorsque le prédicat courant n'est vrai pour aucune paire de régions on passe au suivant et ainsi de suite. Le seuil du prédicat courant est déterminé de manière à ce que les régions créées par les prédicats précédents vérifient celui-ci. L'utilisation de prédicats emboîtés présente de nombreux avantages. Tout d'abord, elle permet de limiter le problème du choix du prédicat. En

effet, il est difficile de choisir un critère pour fusionner les régions d'une image, ce critère étant bien souvent différent pour tous les couples de régions présents. Le choix d'un seul critère introduit donc des erreurs dans la segmentation. L'utilisation de plusieurs prédicats emboîtés permet de remédier à ce problème, ou tout au moins de le limiter. Une difficulté de cette méthode réside néanmoins dans le choix des prédicats et de leurs emboîtements. En effet, quel prédicat favoriser par rapport à quel autre? D'autre part, cette méthode introduit bien souvent un surcoût en temps de calcul par rapport aux méthodes de fusion classiques. Parmi les critères d'uniformité on peut donner:

$$C_{minmax} = \|Max(R_i \cap R_j) - Min(R_i \cap R_j)\| \quad (2.54)$$

$$C_{moy} = \|Moy(R_i) - Moy(R_j)\| \quad (2.55)$$

$$C_{var} = Var(R_i \cup R_j) \quad (2.56)$$

$$C_{fr} = \frac{\sum_{((i,j),(k,l)) \in F(R_i, R_j)} \|I(i, j) - I(k, l)\|}{lf(R_i, R_j)} \quad (2.57)$$

où pour une région  $R_k = \{(x, y)/x = 1, \dots, n, y = 1, \dots, m\}$ , les fonctions *Max*, *Min*, *Moy* et *Var* désignent respectivement le maximum, le minimum, la moyenne, et la variance; elles sont données par:

$$Max(R) = \max_{(i,j) \in R} \{I(i, j)\} \quad (2.58)$$

$$Min(R) = \min_{(i,j) \in R} \{I(i, j)\} \quad (2.59)$$

$$Moy(R) = \frac{1}{card(R)} \sum_{(i,j) \in R} I(i, j) \quad (2.60)$$

$$Var(R) = \frac{1}{card(R)} \sum_{(i,j) \in R} (I(i, j) - Moy(R))^2 \quad (2.61)$$

Dans l'équation (2.57),  $F(R_i, R_j)$  représente la frontière entre les régions  $R_i$  et  $R_j$ , elle est définie par l'ensemble des points  $(p, q)$  tels que:

$$p \in R_i, \quad q \in R_j \quad \text{et} \quad d(p, q) = 1 \quad (2.62)$$

avec  $d(., .)$  une distance qui dans le cas d'un système de 4-voisinage est donnée par [86]:

$$d_4(p, q) = |i - k| + |j - l| \quad \text{où} \quad p = (i, j) \quad \text{et} \quad q = (k, l) \quad (2.63)$$

$lf(R_i, R_j)$  est la longueur de la frontière et est égale à  $card\{F(R_i, R_j)\}$ .

L'opérateur *card* désigne la cardinalité d'un ensemble.

On peut également définir des critères combinant les critères simples précédents. Par exemple:

$$C = a C_{minmax} + b C_{moy} \quad (2.64)$$

$$C = C_{moy}/C_{var}, \quad (2.65)$$

ou plus généralement:

$$C = \sum_i \alpha_i C_i \quad (2.66)$$

où  $a, b, \alpha_i$  sont des poids dont la détermination n'est pas aisée et risque de produire des stratégies ad hoc, impossibles à généraliser. L'utilisation d'un grand nombre de prédicats pour segmenter une catégorie d'images donnée, peut donc conduire à une interdiction de toute généralisation à d'autres types d'images.

### 2.2.1.3.3 Approches par division

Une autre approche pour réaliser la segmentation d'images est celle qui procède par partitionnement. Cette méthode consiste à diviser l'image, qui constitue la région initiale, en régions de plus en plus homogènes. Le processus est réitéré pour chacune des régions produites jusqu'à ce qu'une certaine homogénéité soit atteinte. L'homogénéité d'une région est souvent contrôlée par sa variance ou son contraste. Ces techniques à caractère descendant ont une faiblesse liée à la nature souvent régulière du découpage. Une région est divisée en sous-régions de niveaux inférieurs, les frontières d'une région sont alors représentées sur différents niveaux. Leurs délimitations exactes sont ainsi difficiles à obtenir.

Beaucoup d'algorithmes de division reposent sur l'utilisation des histogrammes de niveaux de gris [251]. Les régions sont alors définies à partir des intervalles entre les vallées.

### 2.2.1.3.4 Approches par division-fusion

Ces méthodes combinent les deux approches précédentes: la division qui partitionne l'image en zones localement homogènes, puis la fusion des régions similaires au sens d'un prédicat de regroupement. Ces deux opérations sont répétées jusqu'à ce qu'elles ne soient plus possibles.

Plusieurs variantes de cette méthode existent. La méthode proposée par Horowitz et Pavlidis [165] sert usuellement de référence. L'image est représentée par un arbre quaternaire (*quadtree*) [304] (Fig.2.2) où chacun de ses noeuds et feuilles représente une région.

Une région est divisée en quatre sous-régions si elle n'est pas homogène. Quatre régions contiguës sont regroupées si elles sont similaires.

FIG. 2.2 - Exemple d'une décomposition spatiale et "quadtree" correspondant.

Pour évaluer l'homogénéité d'une région, Chen et Pavlidis suggèrent l'utilisation d'un test statistique [258]. Ce test suppose que la moyenne ne varie pas trop au cours de la division et que les variances soient connues et constantes.

Soient  $k$  le nombre de pixels de chacune des sous-régions,  $X_{ij}$  le  $j$ ème pixel de la  $i$ ème sous-région,  $\mu_i$  la moyenne de la  $i$ ème sous-région et  $\mu_0$  la moyenne de région avant division. La région est considérée comme homogène lorsque

$$|\mu_i - \mu_0| < \text{seuil}, \quad \forall i \in \{1, 2, 3, 4\} \quad (2.67)$$

En supposant que la région suit une distribution gaussienne, un test permet de vérifier l'hypothèse de l'égalité des moyennes et des variances des sous-régions. Ce test est donné par

$$F = \frac{k \sum_{i=1}^4 (\mu_i - \mu_0)^2 / 3}{\sum_{i=1}^4 \sum_{j=1}^k (X_{ij} - \mu_i)^2 / 4(k-1)} \quad (2.68)$$

Pour de faibles valeurs de  $F$ , la région est considérée homogène.

Dans [31], Brice et Fennema proposent un algorithme composé de deux étapes. Dans la première, l'image est partitionnée en régions de mêmes intensités, et en seconde étape des régions adjacentes sont regroupées sur la base de deux heuristiques utilisant des informations sur les frontières communes et les tailles des régions.

Une autre méthode, conçue dans le but du codage d'images [194], utilise une approximation polynômiale des luminances des régions. L'image est subdivisée en un certain nombre de sous-images. On cherche à approcher chaque sous-image par un polynôme. L'erreur d'approximation est comparée à un seuil prédéfini. Si celui-ci n'est pas dépassé, la sous-image est représentée par les coefficients du polynôme approximant. Sinon, la sous-image est subdivisée en quatre sous-images de demi-taille, et la procédure est réitérée pour chacune de ces nouvelles sous-images jusqu'à ce que toutes les erreurs d'approximation

soient inférieures au seuil fixé.

Après la division, une procédure de fusion de sous-images est effectuée en utilisant un critère de similitude des coefficients de leurs polynômes approximatifs.

Un autre exemple de division-fusion est l'algorithme développé par Beveridge et al. [36]. L'image est préalablement découpée en fenêtres rectangulaires se chevauchant, afin d'éviter tous les problèmes dus aux frontières. La division est ensuite appliquée indépendamment de chacune de ces fenêtres en utilisant les modes et les vallées de leur histogramme de niveaux de gris: une classe est créée pour chaque mode de l'histogramme. Tous les pixels connexes appartenant à la même classe (situés entre les deux vallées d'un pic) sont alors regroupés pour former une région. Une seule itération est donc nécessaire pour la division. La seconde étape (fusion) regroupe les régions suffisamment proches au sens d'un critère appelé "merge-score". Ce critère dépend de la moyenne des niveaux de gris des deux régions, de leur taille et de la longueur de leur contour commun. L'inconvénient de cet algorithme est qu'il introduit beaucoup de seuils: trois seuils sont nécessaires pour extraire les pics des histogrammes locaux, et un seuil est requis pour la fusion.

Nous pouvons également citer l'algorithme développé par Chassery et Melkemi [85]. Cet algorithme, utilisant les polygones de Voronoï, se décompose en trois phases:

- initialisation: elle est réalisée en choisissant des "germes initiaux" autour desquels des polygones de Voronoï sont déterminés. Chaque polygone est caractérisé par la moyenne des niveaux de gris de ses pixels;
- phase de division: de nouveaux noyaux sont introduits de façon itérative dans les polygones non homogènes, modifiant ainsi le diagramme de Voronoï. Cette étape converge lorsque tous les polygones sont homogènes;
- phase de fusion: les polygones similaires sont fusionnés. Ici aussi, un seuil est requis pour la fusion.

Cette méthode présente l'avantage d'utiliser un découpage irrégulier de l'image plus adapté que la structure rigide des *quadtree*. De plus, les différentes étapes de la division peuvent être activées en parallèle.

#### 2.2.1.4 Conclusion

Le sujet abordé dans cette section est celui de la segmentation région par croissance de régions. De manière générale, nous pouvons tirer les conclusions suivantes:

- un inconvénient des méthodes de fusion est qu'elles sont coûteuses en temps de calcul. En effet, la fusion implique, à chaque itération, une recherche sur toute l'image des deux régions les plus aptes à fusionner (celles qui sont les plus similaires). Des méthodes permettant de réduire les temps de calcul existent dans la littérature. Elles consistent

à rechercher des minima locaux et à fusionner les régions correspondant à ces extrema. Malheureusement, la segmentation dépend alors de l'ordre dans lequel les différentes régions sont fusionnées [36].

- un inconvénient des méthodes de division est qu'elles reposent sur des statistiques globales, une petite région située à côté d'une grande sera donc rarement restituée. De plus, de par le principe même de la division, les images sont souvent sur-segmentées [167].
- la division-fusion permet de garantir une segmentation qui vérifie les propriétés d'homogénéité imposées, néanmoins elles sont peu robustes et sensibles aux conditions de départ (la partition initiale, le ou les seuils de tolérance de fusion et/ou division ainsi que le sens du parcours de l'image). Une légère modification de l'un de ces facteurs peut conduire à des segmentations très différentes.

Pour améliorer le résultat de cette méthode et pallier son défaut de robustesse, une solution a été proposée en retenant toutes les fusions possibles à chaque étape et n'autorisant que les meilleures [363].

Notons que les méthodes de division-fusion permettent aussi de réduire le coût algorithmique de la segmentation: plusieurs régions n'ont pas à être fusionnées puisqu'elles n'ont pas été divisées. La fusion ayant un fort coût de calcul par rapport à la division, il est judicieux de ne pas commencer le traitement avec des régions initiales réduites aux pixels. Enfin, un autre avantage de cette méthode de division-fusion est qu'elle conduit à des algorithmes facilement parallélisables: toute une partie du traitement (la division) peut être réalisée de manière séparée et peut donc être partagée sur différents processeurs.

## 2.2.2 Détection de contours

### 2.2.2.1 Introduction

La détermination des limites des objets dans une scène présente un intérêt primordial pour le traitement de l'image. La détection de contours est ainsi un sujet de recherche très important en traitement d'images, et les contours ont été et sont toujours beaucoup utilisés comme primitives en analyse d'images et en vision par ordinateur. La qualité et la précision des contours détectés jouent donc un rôle très important chaque fois que l'on doit mettre en correspondance des primitives robustes issues d'images différentes. Citons par exemple: la stéréovision, la vision active, l'analyse du mouvement, la reconnaissance des formes, etc...

Contrairement aux approches de segmentation par extraction de régions qui recherchent les zones de l'image qui vérifient des propriétés d'homogénéité au sens d'un ou de plusieurs critères donnés, les méthodes de segmentation par détection de contours recherchent des discontinués des intensités lumineuses ou des ruptures de modèles de textures.

Le problème de la détection de contours dans les images a été largement abordé ces deux dernières décennies et il existe une littérature abondante sur ce sujet [17] [60] [103] [108] [171] [231] [269] [274] [283] [284] [294] [306] [332] .

Les contours se manifestent dans l'image par une forte transition des valeurs d'intensité. La détection de ces variations s'opèrent généralement selon deux catégories. Dans la première catégorie se trouvent les techniques qui emploient un filtrage de type différentiel ou adaptatif. Elles sont fondées sur la recherche des points de l'image présentant un fort gradient, une forte courbure ou encore une forte corrélation avec un profil prédéfini. Les techniques qui n'utilisent pas le filtrage sont essentiellement basées sur l'utilisation des outils de la programmation dynamique, la morphologie mathématique, les approches probabilistes et autres.

Notons que d'une manière générale, chaque méthode de détection est justifiée par la définition de type de contour recherché. La plupart des méthodes de détection de contours sont basées sur la recherche de des modèles de contours de type "marche" additionné à un bruit blanc.

Nous décrivons dans les paragraphes qui suivent quelques unes des ces méthodes et discutons les limites et les avantages de chacune d'elles.

### 2.2.2.2 Opérateurs du premier et du second ordre

Parmi les opérateurs du premier ordre les plus courants on peut citer ceux de Prewitt et de Sobel [284] [316], l'opérateur de Kirsh [193] et l'opérateur de Nagdif. Ce dernier proposé par Cocquerez et Devars dans [69] est plus complexe car il effectue, tout d'abord, un lissage non linéaire de l'image en utilisant un filtrage proposé par Nagao [248], suivi d'une dérivation réalisée par un opérateur gradient. Cet opérateur permet d'une part de diminuer le bruit, d'autre part d'amplifier les contrastes entre différents objets.

Dans la catégorie d'opérateurs du second ordre, les plus utilisés sont: l'opérateur laplacien et l'opérateur laplacien-gaussien [231].

Les contours détectés par les techniques exploitant le laplacien sont relativement de meilleure qualité et sont de largeur d'un point. L'opérateur laplacien-gaussien (LoG) est en plus insensible aux petites variations non significatives.

Notons que tous les opérateurs cités sont intéressants et donnent des résultats satisfaisants dans le cas d'images simples. Ils sont néanmoins parfois moins précis pour la localisation des contours. Les résultats obtenus par la plupart d'entre eux sur des images complexes restent très sensibles au bruit et engendrent souvent de fausses détections.

### 2.2.2.3 Approches utilisant une modélisation surfacique

Les deux approches présentées dans ce qui suit opèrent par une approximation polynomiale sur l'image intensité ou l'image du laplacien. Une détection analytique est alors effectuée sur la fonction approximante. L'avantage principal est de pouvoir disposer de

contours extraits à une résolution inférieure au pixel “subpixellique”.

La première approche proposée par Huertas et Medioni dans [162] permet l'extraction de contours avec une localisation subpixellique à l'aide d'un modèle surfacique sur l'image LoG. L'hypothèse de base est qu'au voisinage d'un passage par zéro, l'image laplacien peut être approximée à l'aide d'un polynôme de degré 3 sur les lignes et les colonnes. Les passages par zéro identifiant les contours sont alors détectés de manière analytique à l'aide du polynôme approximant et non plus sur l'image laplacien. Les fonctions d'interpolation utilisées sont les polynômes de Chebyshev dans un voisinage de taille 3 x 3 centré sur un passage par zéro détecté sur l'image laplacien.

Dans une seconde approche proposée par Haralick [154], l'image est assimilée à une surface échantillonnée dont la troisième dimension est la valeur des niveaux de gris (appelée modèle de facette). L'auteur propose de caractériser l'intensité dans un voisinage d'un point de l'image par un polynôme de second degré à deux variables. Les contours sont situés sur les plus fortes pentes de la facette, ce qui correspond aux passages par zéro d'une dérivée seconde en présence d'un gradient non nul.

L'avantage de cette technique réside dans la modélisation de la facette qui ne présente pas de pente véritable en présence de bruit. Elle est néanmoins assez lourde en temps de calcul et nécessite un grand nombre de points ce qui limite son efficacité et son utilisation. Pour surmonter ces difficultés de robustesse aux bruits, de précision de localisation et d'efficacité de détection, des approches par optimisation sont proposées.

#### 2.2.2.4 Approches par optimisation

Parmi les approches par optimisation, celles développées par Canny [60], Deriche [108] et Shen et Castan [308] semblent être les plus intéressantes et méritent qu'on s'y attarde.

##### 2.2.2.4.1 Filtre de Canny

L'approche développée par Canny [60] a consisté en la recherche de l'opérateur de dérivation optimal sous forme de filtre à réponse impulsionnelle finie (RIF). Cette approche peut être résumée comme suit.

Soit  $I(x)$  un signal monodimensionnel représentant un saut d'amplitude  $A$ , noyé dans un bruit gaussien  $n(x)$  de moyenne nulle et de variance  $\sigma_0^2$ :

$$I(x) = Au_{-1}(x) + n(x) \quad (2.69)$$

où  $u_{-1}$  désigne la fonction de Heaviside.

Soit  $\Theta(x_0)$ , la sortie au point  $x_0$  de la convolution du signal  $I(x)$  avec un opérateur de détection  $f(x)$ :

$$\Theta(x_0) = \int_{-\infty}^{+\infty} I(x)f(x_0 - x)dx \quad (2.70)$$

Le problème posé est de trouver  $f(x)$  tel que  $\Theta(x_0)$  soit maximum sous les 3 contraintes suivantes :

1. *Bonne détection*: ce critère revient à chercher  $f(x)$  antisymétrique et tel que le rapport signal sur bruit  $S/B$  soit maximum:

$$S/B = \frac{A|\int_{-W}^0 f(x)dx|}{\sigma_0\sqrt{\int_{-W}^{+W} f^2(x)dx}} \quad (2.71)$$

2. *Bonne localisation*: ce critère correspond à la minimisation de la variance  $\sigma^2$  de la position des passages par zéro et revient à maximiser la localisation  $\Lambda$  définie comme l'inverse de  $\sigma^2$ :

$$\Lambda = \frac{A|f'(0)|}{\sigma_0\sqrt{\int_{-W}^{+W} f'^2(x)dx}} \quad (2.72)$$

3. *Non multiplicité des réponses*: ce critère correspond à la limitation du nombre de maxima locaux détectés en réponse à un seul contour. La distance moyenne entre les maxima locaux , notée  $x_{max}$  est alors contrainte à l'égalité suivante:

$$x_{max} = 2\pi \left( \sqrt{\frac{\int_{-W}^{+W} f'^2 dx}{\int_{-W}^{+W} f''^2(x) dx}} \right) \quad (2.73)$$

Trouver  $f(x)$  qui maximise le produit  $S/B.\Lambda$  sous la contrainte que le 3ème critère soit fixé à une constante  $k$  revient alors à trouver la solution de l'équation différentielle suivante :

$$2f(x) - 2\lambda_1 f''(x) + 2\lambda_2 f''''(x) + \lambda_3 = 0 \quad (2.74)$$

qui admet comme solution générale de la forme:

$$f(x) = a_1 e^{\alpha.x} \sin(\omega.x) + a_2 e^{\alpha.x} \cos(\omega.x) + a_3 e^{-\alpha.x} \sin(\omega.x) + a_4 e^{-\alpha.x} \cos(\omega.x) \quad (2.75)$$

Afin de déterminer l'opérateur  $f(x)$  sous forme de filtre à réponse impulsionnelle finie (RIF) définie sur l'intervalle  $[-W, +W]$  et présentant une pente  $S$  à l'origine, Canny a imposé les conditions aux limites suivantes :

$$f(0) = 0, f(W) = 0, f'(0) = S, f'(W) = 0 \quad (2.76)$$

Ces 4 conditions aux limites permettent alors aux coefficients  $a_1...a_4$  d'être déterminés.  $f(x)$  étant impaire, la solution est étendue aux  $x$  négatifs avec  $f(x) = -f(-x)$ .

En utilisant une optimisation numérique sous contrainte, Canny trouva que l'opérateur le plus performant correspondait à un indice de performance  $S/B.\Lambda = 1.12$ . L'opérateur n'étant pas simple à mettre en œuvre, il l'approxima au vu de sa forme par l'opérateur première dérivée d'une gaussienne qui présente un indice de performance  $S/B.\Lambda = 0.92$  et  $k = 0.51$ .

### 2.2.2.4.2 Filtre de Deriche

L'approche développée par Deriche [108] est similaire à celle de Canny mais l'opérateur optimal dans ce cas est mis sous forme d'un filtre à réponse impulsionnelle infinie (RII). Il en résulte une équation différentielle admettant comme solution l'opérateur suivant:

$$f(x) = \frac{S}{\omega} e^{-\alpha|x|} \sin(\omega x) \quad (2.77)$$

Évaluant les différentes intégrales intervenant dans le calcul des indices de performance de cet opérateur, les résultats suivants sont obtenus:

$$\Lambda = \sqrt{2\alpha} \quad S/B = \sqrt{\frac{2\alpha}{\alpha^2 + \omega^2}} \quad (2.78)$$

$$S/B.\Lambda = \frac{2\alpha}{\sqrt{\alpha^2 + \omega^2}} \quad k = \sqrt{\frac{\alpha^2 + \omega^2}{5\alpha^2 + \omega^2}} \quad (2.79)$$

En posant  $\alpha = m\omega$ , il obtient les 3 cas suivants:

$$\begin{aligned} 1. \quad m \gg 1 \quad \Lambda &= \sqrt{2\alpha} \quad S/B = \sqrt{2/\alpha} \quad S/B.\Lambda = 2 \quad k = 0.44 \\ 2. \quad m = 1 \quad \Lambda &= \sqrt{2\alpha} \quad S/B = \sqrt{1/\alpha} \quad S/B.\Lambda = \sqrt{2} \quad k = 0.58 \\ 3. \quad m = \sqrt{3} \quad \Lambda &= \sqrt{2\alpha} \quad S/B = \sqrt{3/2\alpha} \quad S/B.\Lambda = \sqrt{3} \quad k = 0.5 \end{aligned} \quad (2.80)$$

Le cas 3 montre que pour une valeur identique de  $k$ , l'indice de performance de l'opérateur Deriche est meilleur que la première dérivée d'une gaussienne.

Le cas 2 montre que pour une valeur identique de  $k$ , l'opérateur Deriche présente un indice de performance amélioré de près de 25% par rapport à l'opérateur de Canny  $S/B.\Lambda = 1.12$  et  $k = 0.58$ .

Le cas 1 présente le meilleur indice. Il correspond à la limite de l'opérateur Deriche pour  $\omega$  tendant vers zéro. Il est facile de vérifier que cette limite correspond à l'opérateur  $f(x)$  donnée par :

$$f(x) = Sx e^{-\alpha|x|} \quad (2.81)$$

En fait cet opérateur correspond tout simplement à la solution de l'équation différentielle 2.74 dans le cas où le discriminant de son équation caractéristique est nul. Ses indices de performance correspondent à ceux donnés dans le cas 1.

Le paramètre  $\alpha$  doit être choisi assez grand si on veut favoriser la localisation au détriment de la détection (pour les signaux très peu bruités) et assez petit si on veut privilégier la détection dans le cas de signaux très bruités.

Après l'opération de filtrage, la détection de contours peut s'opérer de deux manières: soit par une extraction des maxima locaux de la dérivée première, soit par l'extraction des passages par zéro de la seconde dérivée. Cette opération est ensuite suivie par un seuillage de type hystérésis dont le principe est de seuiller d'abord avec un seuil bas et d'éliminer ensuite toutes les chaînes n'ayant pas au moins un point dont la norme du gradient est supérieure au seuil haut. L'avantage de la procédure de seuillage par hystérésis est qu'elle permet d'obtenir plus d'ensembles de points connexes, ce qui facilite grandement les étapes de chaînage et d'approximation prévues généralement en amont de cette chaîne.

L'extraction suivant la technique de filtrage optimal au sens des critères de Canny-Deriche, est décrite en détail dans l'annexe A.

### 2.2.2.4.3 Filtre de Shen et Castan

En s'inspirant du formalisme mathématique développé par Canny, l'approche proposée par Shen et Castan [101] [308] a consisté en l'élaboration d'un filtre passe bas optimal à partir duquel est déduit un filtre passe haut pour la détection des contours. Un modèle de saut d'amplitude bruité est considéré et un critère à optimiser est alors dérivé. Il ressort de cette étude que le filtre linéaire optimal de lissage recherché est le filtre exponentiel symétrique:

$$f(n) = ae^{-\alpha|n|} \quad (2.82)$$

Si  $y(n)$  correspond à la réponse d'un système ayant  $f(n)$  comme réponse impulsionnelle et  $x(m)$  comme entrée d'un système, Shen et Castan proposent de le mettre en œuvre comme suit:

$$y_1(n) = x(n-1) + e^{-\alpha}y_1(n-1) \quad \text{pour } n = 1, \dots, M \quad (2.83)$$

$$y(n) = y_1(n+1) + e^{-\alpha}y(n+1) \quad \text{pour } n = M, \dots, 1 \quad (2.84)$$

Etant donnée une image  $I$ , elle est d'abord filtrée par un opérateur de lissage exponentiel. Le résultat est alors soustrait de l'image originale pour approximer le laplacien. Une image binaire est alors produite en mettant à 1 les valeurs positives et à 0 le reste des données. Les frontières des zones positives et négatives forment alors les points de contours candidats. Une suppression des points isolés ainsi que les groupes connexes de points ayant une surface inférieure à un seuil donné est alors effectuée. Pour distinguer les vrais points de contours de ceux correspondant à du bruit, un seuillage sur le gradient est ensuite réalisé. Le calcul du gradient se fait suivant une technique adaptative. On calcule pour chaque passage par zéro et à partir de la fenêtre correspondante dans l'image originale les moyennes de niveaux de gris correspondant aux zones 1 et 0 de l'image binaire. Le gradient adaptatif est alors égal à la différence entre ces deux moyennes.

Une image ne contient pas uniquement des discontinuités de type marche. Il existe des applications où la recherche de profils de ce type n'est pas adaptée et où il est nécessaire de prendre en compte d'autres profils de contours.

Des travaux relatifs à l'extraction de contours qui ne sont pas de type "marche" ont été proposés. On trouvera dans [372] une approche similaire de celles de Canny et de Deriche mais appliquée pour la détection de contours de type "pic" ou "toit". La modélisation de ce type de contour aboutit à un opérateur correspondant au laplacien d'une fonction de lissage de type exponentiel. Dans [269] Petrou et al. ont conçu un détecteur optimal de rampe. Perona et Malik [274] proposent d'utiliser une combinaison de filtres quadratiques pour détecter n'importe quel type de discontinuité.

D'autres techniques basées sur des représentations multi-résolution de l'image [26] [336], ont été présentées pour résoudre le problème des distorsions locales et pour obtenir uniquement les contours significatifs.

#### 2.2.2.4.4 Conclusion

L'approche proposée par Canny sert toujours de référence dans la littérature car sa solution est correctement formalisée et le comportement du filtre est ainsi connu avec précision. Toutefois, certaines limites peuvent être recensées dans ce modèle:

- La forme "marche d'escalier" du contour et l'hypothèse du bruit blanc gaussien affectant l'image. Il arrive très souvent que la forme du signal soit beaucoup plus complexe, comme cela a été montré par Perona et Malik dans [274]. Par ailleurs, si l'image est dégradée par un bruit multiplicatif, ce qui est le cas des images obtenues par un système cohérent (laser ou radar), le modèle de Canny est inadapté. Le problème de la dégradation des textures notamment, celles ayant des faibles variations, se pose aussi quand le bruit est relativement faible.
- Le seuillage par hystérésis permet de sélectionner les contours les plus représentatifs. Il faut alors fixer les deux paramètres  $S_h$  et  $S_b$  correspondant respectivement au seuil haut et au seuil bas. Pour cette nouvelle étape, un compromis est, là aussi, nécessaire pour conserver les contours correspondant aux frontières entre différents objets de la scène et les "faux contours" dus au bruit. Par tâtonnement, on trouve les seuils les mieux adaptés à l'image.
- Ce type de filtrage permet de détecter correctement les contours, mais le fait d'augmenter la valeur de  $\alpha$  ne suffit pas pour détecter la totalité des contours. En effet, ils présentent quelques interruptions notamment au voisinage des coins et des "jonctions en T".

- Notons enfin que de nombreux aspects liés à la discrétisation du signal sont occultés lors de la généralisation à deux dimensions d'un filtre "optimal" dans l'espace continu mono-dimensionnel.

Des solutions à certains de ces problèmes ont été récemment proposées par Chehikian dans [76]. L'auteur suggère une nouvelle classe de filtres de lissage, à partir desquels un filtre de dérivation est déduit. La démarche employée est différente de celle de Canny, Deriche, Shen et Castan. Elle vise, partant d'un modèle discret de contour, à lui associer une fonction continue dont la dérivée est l'estimation cherchée du gradient. D'une part, aucune hypothèse sur la nature du bruit n'a été faite, d'autre part, l'opérateur proposé permet une interpolation spline des gradients afin de produire une carte de contours connectés, lisses et ayant une résolution "subpixellique" (de taille double de l'image originale). Dans cette carte les points de contours de coordonnées paires correspondent aux points de l'image d'origine. Les points dont une coordonnée est impaire, s'interprètent comme des arêtes entre points. Les points dont les deux coordonnées sont impaires, s'interprètent comme des sommets entre points. Cette forme de représentation des points de contours permet ainsi de distinguer les points qui constituent les régions et les points de contours qui en constituent les frontières. L'inconvénient majeur de cette approche réside dans le coût algorithmique relativement élevé, nécessaire pour l'interpolation. Cependant, un sous-échantillonnage du gradient interpolé, en ne conservant que les points ayant les coordonnées toutes deux impaires, permet de réduire considérablement le temps de calcul [76].

### 2.2.2.5 Autres approches de détection de contours

Dans les méthodes de détection de contours présentées précédemment, le but était de modéliser le contour à détecter. Ce qui n'est pas le cas d'autres catégories de méthodes, parmi lesquelles on trouve: l'approche variationnelle, l'approche statistique, l'approche basée sur la programmation dynamique, l'approche morphologique,...

Dans l'*approche variationnelle* (contours actifs) [201], les contours sont modélisés par des courbes élastiques qu'il s'agit d'ajuster le mieux possible. Pour cela on associe une énergie dont la définition est étroitement liée à celle des points contours. Les minima locaux de cette énergie correspondent aux contours s'ajustant le mieux au modèle utilisé.

L'inconvénient majeur de cette approche réside dans la phase d'initialisation manuelle de la courbe de départ. Elle présente également une sensibilité au choix des paramètres de rigidité et d'élasticité. Cependant, son intégration avec une approche pyramidale (analyse multi-échelle) permet de s'affranchir de cette faiblesse de robustesse [70].

*Les méthodes statistiques* de détection de contours sont en quelque sorte duales des précédentes par le fait qu'elles se basent sur la modélisation des régions qui séparent les contours à détecter. Les frontières des composantes de l'image sont dans ce cas mises en évidence par des ruptures de modèles aux limites [120]. Chellappa et al. [377] travaillent sur des modèles

linéaires (AR et ARMA) bidimensionnels non stationnaires faisant intervenir des échantillons présents pour prédire les contours suivants.

D'autres travaux ont été réalisés pour la détection de contours en utilisant les modèles markoviens, notamment depuis les travaux de Geman et Geman [138].

Dans le cas de *méthodes fondées sur la programmation dynamique* [224], la détection de contours est traduite par la minimisation d'une fonction coût qui tient compte des caractéristiques du contour recherché (contraste, courbure,...). Cette fonction additive s'obtient en ajoutant à chaque étape des coûts locaux pour représenter un coût global.

Dans [329], Tan, Gelfand et Delp ont proposé de minimiser une fonction de coût pour détecter les contours en utilisant le recuit simulé. Leur algorithme se décompose en deux étapes:

Tout d'abord, une matrice de contours potentiels est créée. A l'aide d'un filtrage suivi d'une suppression des non-maxima locaux, la valeur impulsionnelle du filtre associé à un élément de contour le mieux adapté est conservée. Pour réaliser le filtrage, douze noyaux de convolution, chacun d'eux étant associé à un profil de contour sont utilisés.

Puis, une fonction d'énergie est définie comme une somme pondérée de facteurs de coût.

Dans [310] et [156], *des méthodes utilisant la morphologie mathématique* ont été proposées. Une approximation du module du gradient en un point de l'image est obtenue par une combinaison des opérations de dilatation et d'érosion. L'application de cette méthode conduit généralement à des résultats de sur segmentation. Son avantage réside principalement dans sa simplicité.

### 2.2.2.6 Quelques méthodes de fermeture de contours

Intuitivement, un point contour est défini comme un point de l'image où l'on observe une discontinuité de la fonction qui représente l'image. D'une manière générale, l'extraction des contours, dans le cas des images réelles se heurte à des problèmes qui peuvent engendrer:

- une sur-détection souvent associée à une présence de bruits, une irrégularité du contraste due à une variation des conditions d'éclairage (effets d'ombres, reflets, et autres), ou une forte variation de l'intensité lumineuse causée par la présence de texture.
- une sous-détection souvent liée à des éléments difficiles à extraire tels que des points angulaires ou les points des jonctions de lignes dans le cas d'images d'intérieur.

Malgré toutes les améliorations apportées par les différentes techniques vues précédemment, ce niveau de détection aboutit donc souvent à une image de contours bruitée et fractionnée. Les contours obtenus par un détecteur de contours classique, par exemple le filtre de Canny-Deriche, présentent certaines lacunes, notamment des interruptions au niveau des coins et des "jonctions en T". De plus, il faut trouver un compromis pour maximiser le rapport signal sur bruit. Des traitements et procédés supplémentaires qui opèrent directement sur

ces contours en tenant compte du contexte sont requis. Ils ont pour but d'éliminer les bruits dus aux effets de la sur-détection, ou prolonger les fragments et les connecter dans le cas d'une sous-détection.

Plusieurs approches ont été proposées [93] pour la fermeture de contours, celle-ci représentant une étape fondamentale pour l'interprétation de scènes: exploration de graphe, groupement perceptuel, utilisation des contours déformables et modélisation markovienne.

### Recherche du meilleur chemin à partir d'une extrémité

Deriche et al. [109] ont proposé une méthode permettant de compléter les interruptions de contours. Le but est de trouver, à partir de chaque extrémité de contour, le chemin suivant la meilleure ligne de crête dans l'image de la norme du gradient. Le principe de l'algorithme de fermeture de contours est le suivant. Un balayage de l'image de contours est effectué ligne par ligne, du haut vers le bas. Quand une extrémité est rencontrée, une procédure permettant de trouver le meilleur candidat pour fermer les contours est activée. Quand la condition d'arrêt est atteinte, le balayage de l'image continue à partir de l'endroit où il s'était interrompu.

### Groupement perceptuel

Les études neurobiologiques mettent en évidence le rôle important de l'organisation perceptuelle des contours dans le processus visuel. Le problème est d'utiliser certaines propriétés de la vision naturelle pour élaborer un algorithme de groupement de contours.

Plusieurs approches ont été proposées pour traiter le problème du groupement perceptuel.

Lowe [216] a proposé une théorie de l'organisation perceptuelle s'appuyant sur le principe de la non-accidentalité. Ce modèle permet d'attribuer un degré de signification à chaque groupement possible d'une scène donnée et de conserver uniquement les groupements les plus significatifs.

Grossberg et Mingolla [147] ont élaboré un modèle connexionniste complexe de perception visuelle bas niveau. Ce travail utilise l'existence de contours virtuels et repose sur une approche globale de la perception visuelle. Ils utilisent différentes caractéristiques telles que la texture, les contours et les intensités lumineuses.

Gideon et Medioni [146] ont utilisé une méthode de groupement perceptuel décrite en fonction des propriétés suivantes: la cocircularité entre segments appartenant à la même courbe, la constance de la courbure, la préférence pour les faibles courbures et la proximité des segments engendrant le contour.

Citons, enfin, Parent et Zucker [288] qui proposent une approche prenant en compte des contraintes de position, de direction et de courbure. Ces contraintes leur permettent d'évaluer la compatibilité entre deux segments donnés, en utilisant la notion de segments cocirculaires.

### 2.2.2.7 Conclusion

Quelques approches de base d'extraction de contours ont été présentées. A partir de cette présentation et de l'étude comparative menée dans [94], sur les différentes méthodes d'extraction de contours, nous pouvons tirer les conclusions suivantes.

- Les transitions franches, dans un milieu peu bruité, sont bien extraites par tous les traitements, l'opérateur nagdif ayant l'avantage de la rapidité d'exécution.
- Dans le cas de transitions lentes, correspondant soit à des contours "flous", soit à des variations d'éclairement, on observe un mauvais comportement de nagdif, qui va créer des contours hachés ou doublés. Par contre, comme il est possible de régler la taille des filtres de Canny-Deriche par le choix du paramètre  $\alpha$ , cet opérateur est mieux adapté à ce type de transition.
- Les détecteurs optimaux tendent ainsi à trouver le meilleur compromis entre une bonne détection et une bonne localisation des contours. De même, le seuillage effectué sur le gradient ou le laplacien ne produit jamais des contours fermés, car il y a également un compromis entre la détection de vrais contours et celle des points dus au bruit. Les opérations de fermeture de contours utilisées ne semblent pas toujours résoudre ce problème de façon satisfaisante.
- Les contours de très faible amplitude, même en milieu bruité, sont particulièrement bien détectés par le filtrage de Canny-Deriche.
- Les contours de zones texturées sont mal extraits par nagdif, mais peuvent l'être par Canny-Deriche, pourvu que les intensités moyennes soient différentes de part et d'autre des contours.
- Les méthodes basées sur l'utilisation du gradient ou de la dérivée première d'une fonction de lissage sont mal adaptées à l'extraction de points anguleux.
- Les opérateurs utilisant la dérivée seconde sont plus sensibles au bruit et donnent des contours moins "lisses" que ceux fournis par les opérateurs employant le gradient.

A partir de ces remarques, la technique de Canny-Deriche nous semble la plus intéressante. Notons, toutefois, que les contours obtenus par ce type de filtrage ne sont pas tous significatifs.

D'un point de vue général, analyser une scène suppose dans un premier temps d'en isoler les éléments constitutifs. L'information donnée par les contours n'est pas souvent pertinente pour les différentes applications d'analyse d'images. En effet, des objets distincts possèdent le plus souvent des aspects de surface différents par leur couleur, leur texture, ou leurs conditions d'éclairement. Ces contrastes forment les contours de l'image; la correspondance entre limite physique des objets et contours de l'image n'est pas absolue. Deux surfaces de textures différentes peuvent facilement être distinguées sans avoir recours à leurs contours.

Les contours les plus intéressants sont les limites des objets ou des surfaces de la scène, l'intégration de la texture est alors nécessaire pour corriger les contours indésirables se trouvant à l'intérieur de ce type de surface.

La segmentation d'images par analyse de textures fait l'objet de la section qui suit.

### 2.2.3 Extraction de régions texturées

#### 2.2.3.1 Définition et présentation du problème

Dans un souci de segmenter un large éventail de type d'images, la prise en compte du paramètre "texture" s'avère indispensable. En effet, beaucoup d'images, et plus particulièrement les images satellitaires, les images médicales et les images de scènes d'extérieur, possèdent des zones texturées, non homogènes au sens des niveaux de gris (figure 2.3). Nous voyons donc la nécessité d'introduire des paramètres de texture dans la méthode de segmentation.

Malgré la diversité et le nombre important des travaux consacrés à l'étude de texture [27] [137] [153] [180] [264] [79] [82] [84] [145], la définition précise de celle-ci n'est toujours pas donnée. La difficulté principale provient du fait que la texture est étroitement liée à la perception visuelle humaine et par conséquent contient une composante subjective se prêtant mal à une formalisation.

La définition la plus répandue détermine la texture comme étant une région de l'image dont l'observation se traduit par une impression visuelle d'homogénéité pour toutes les transitions possibles à l'intérieur de cette région. Elle traduit une répartition spatiale d'un même motif dans différentes directions de l'espace. Ceci lui attache donc deux dimensions sur la base desquelles on peut la décrire. La première concerne les motifs élémentaires constituant la texture. La seconde implique la distribution spatiale de ces motifs [155].

Un motif élémentaire est un ensemble connexe de points caractérisé par des propriétés de nature quantitative telles que les moments, l'entropie et l'énergie entre autres et par des propriétés géométriques comme sa forme, sa surface, etc...

La répartition spatiale des motifs nécessite la définition d'une fonction permettant la prise en compte des intercorrélations ou d'un ensemble de règles de placement des motifs sur le support bidimensionnel.

Les approches les plus fréquemment rencontrées dans la littérature pour l'analyse de texture peuvent être classées en deux grandes classes: l'approche statistique et l'approche structurelle.

- *L'approche statistique* repose sur l'analyse des statistiques des zones plus ou moins

étendues de la texture. Elle doit être employée dans le cadre d'une vision microscopique des textures, et s'applique aux textures à caractère anarchique mais homogène. Les méthodes de cette approche considèrent la texture comme la réalisation d'un processus stochastique bidimensionnel. Elles se prêtent à l'étude des textures à caractère aléatoire où il est difficile d'isoler des motifs élémentaires. Deux catégories de méthodes peuvent être distinguées dans cette approche: celle basée sur l'extraction de paramètres puis classification et celle basée sur l'adaptation d'un modèle.

- *L'approche structurelle* considère les textures comme un agencement spatial d'un ensemble de motifs de base suivant certaines règles de placement. Cette approche qui cherche à mettre en évidence une structuration plus ou moins régulière des éléments de la texture et correspond à une vision plutôt macroscopique de celle-ci. Elle est particulièrement adaptée au cas où les motifs sont bien discernables.

FIG. 2.3 - *Echantillons de textures pris de l'album de Brodatz.*

Il existe toutefois des méthodes qui combinent les deux approches en tenant compte à la fois de l'aspect micro et macroscopique de certaines textures. Par exemple, dans le cas de

la texture des sillons d'un champ de végétation, l'arrangement des sillons sera caractérisé par une approche structurale alors que l'état de surface de la végétation sera modélisé par une approche statistique. Ces observations sous entendent une notion importante qui est celle de l'existence d'une échelle de résolution à laquelle les textures peuvent être analysées d'une façon optimale. L'aspect microscopique correspond aux hautes résolutions et l'aspect macroscopique aux basses résolutions.

Les domaines d'application de l'approche structurale étant relativement limités, nous ne présenterons dans les paragraphes suivants que les principales méthodes des deux catégories de l'approche statistique.

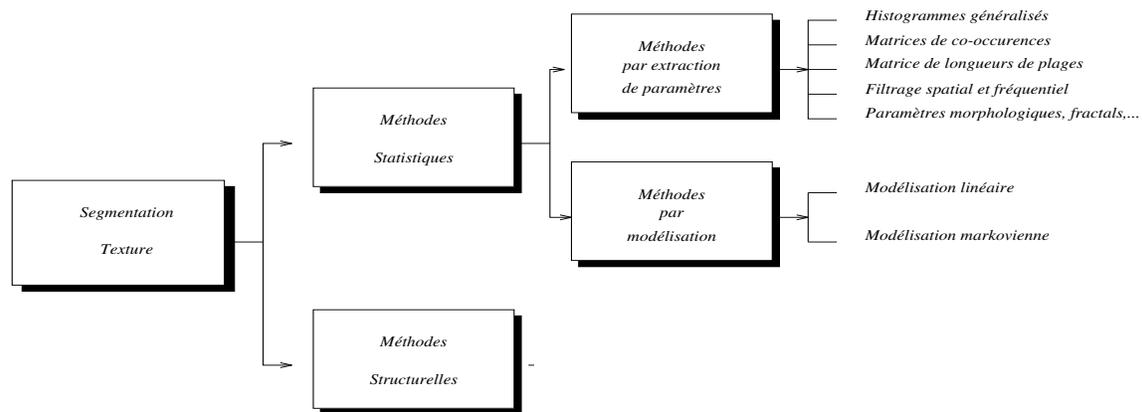


FIG. 2.4 - Classification des méthodes de segmentation de textures

### 2.2.3.2 Approches par extraction de paramètres et classification

On s'intéresse ici à l'extraction de mesures statistiques par analyse d'ensembles de points appartenant à une fenêtre d'observation. Ces paramètres peuvent ensuite être utilisés comme attributs en entrée de schémas de classification et d'identification.

Il s'agit ainsi de définir un vecteur d'attributs dont les composantes doivent être représentatives, pertinentes et discriminantes de façon qu'il puisse identifier sans ambiguïté une texture parmi d'autres. Les paramètres sont donc déterminés suite à une analyse qui peut être faite de manière globale ou de manière locale. L'analyse peut se faire soit dans le domaine spatial soit dans le domaine fréquentiel et de façon adaptative ou non.

L'extraction de paramètres de texture a fait l'objet de nombreux travaux, parmi lesquels on peut citer les méthodes d'analyse basées sur l'utilisation de statistiques du premier et du second ordre qui englobent entre autres les histogrammes de niveaux de gris, les matrices de co-occurrences, les longueurs de plages, les dépendances spatiales, et les méthodes fondées sur l'analyse par filtrage spatial ou fréquentiel.

### 2.2.3.2.1 Histogramme généralisé des niveaux de gris

Soit  $\Delta = \{\vec{\delta}_0, \vec{\delta}_1, \vec{\delta}_2, \dots, \vec{\delta}_{N-1}\}$  un ensemble de vecteurs défini dans le plan.

On définit  $\vec{\delta}_k$  comme étant le  $k$ ième vecteur de déplacement et  $N$  dénote l'ordre du système des vecteurs dans le plan.

La figure 2.5 donne une illustration de l'ensemble  $\Delta$  dans le cas d'ordre  $N$ .

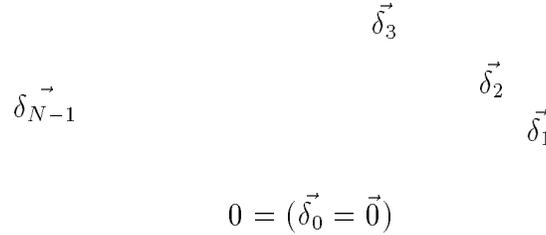


FIG. 2.5 - Vecteurs de déplacement.

L'histogramme  $h_N$  d'ordre  $N$  relatif à  $\Delta$  est défini par:

$$h_N(g_0, g_1, \dots, g_{N-1}; \vec{\delta}_0, \vec{\delta}_1, \dots, \vec{\delta}_{N-1}) \equiv \text{Prob}(f_{\vec{\delta}_0} = g_0, f_{\vec{\delta}_1} = g_1, f_{\vec{\delta}_2} = g_2, \dots, f_{\vec{\delta}_{N-1}} = g_{N-1}) = \frac{N_g}{N_T} \quad (2.85)$$

où

$f_{\vec{\delta}}$  représente la valeur du niveau de gris du pixel avec un déplacement  $\vec{\delta}$  de l'origine.

$N_g$  correspond au nombre de groupes de pixels  $\{f_{\vec{\delta}_0} = g_0, f_{\vec{\delta}_1} = g_1, f_{\vec{\delta}_2} = g_2, \dots, f_{\vec{\delta}_{N-1}} = g_{N-1}\}$

$N_T$  correspond au nombre total de groupes de  $N$  pixels dans l'image.

$h_N$  correspond à un histogramme multi-dimensionnel. Un élément de cet ensemble représente le nombre (normalisé) de fois où un  $N$ -uplet de niveaux de gris  $g_0, g_1, \dots, g_{N-1}$  situés à des distances  $\delta_j$  ( $j \in \{0, 1, \dots, N-1\}$ ) de l'origine, apparaissent dans une zone de l'image.

En utilisant cette définition, on peut déterminer:

- l'histogramme du premier ordre de l'image ( $h(g) = h_1(g_0; \vec{\delta}_0)$ );
- l'histogramme du second ordre ( $h_2(g_0, g_1; \vec{\delta}_1)$  où  $\vec{\delta}_1 = (d, \theta)$ ) qui est égal à la matrice de co-occurrence  $p(g_0, g_1; d, \theta)$  après normalisation;
- et les histogrammes d'ordres supérieurs.

### 2.2.3.2.2 Matrices de co-occurrences

La méthode de la matrice de co-occurrences ou de l'histogramme du second ordre, développée par Haralick et al. [168], pour la discrimination des textures consiste à estimer les fréquences d'apparition de couples de niveaux de gris de pixels dans une zone de l'image. Les matrices de co-occurrence sont fonctions de  $d$  et  $\theta$  (vecteur de déplacement), on peut calculer autant de matrices qu'il y a de distances  $d$  et d'angles  $\theta$  possibles pour une image. Généralement on les établit pour des valeurs de  $\theta = k\pi/4$  ( $k = 0, 1, 2, 3$ ) et des valeurs de  $d = 1, 2$ .

Pour normaliser les données de la matrice et leur donner une signification en termes de probabilités d'occurrence, chacun des ses éléments est divisé par la somme de tous les éléments de la matrice.

#### • Attributs de texture

Les matrices de co-occurrences contiennent une grande quantité d'informations, difficilement exploitables dans leur intégralité [133]. Il faudra donc réduire cette quantité d'informations et garder des données pertinentes, relatives à la texture de la scène.

De ces matrices sont donc extraites des mesures qui sont porteuses d'informations d'uniformité, de contraste, d'homogénéité, de complexité de transition, de structure de texture. Nous en retiendrons les principaux attributs suivants.

L'énergie

$$EN_{d,\theta} = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (p(i, j, d, \theta))^2 \quad (2.86)$$

Le contraste

$$CT_{d,\theta} = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i - j)^2 p(i, j, d, \theta) \quad (2.87)$$

La corrélation

$$CR_{d,\theta} = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} \frac{(i - \mu_x)(j - \mu_y)p(i, j, d, \theta)}{\sigma_x \sigma_y} \quad (2.88)$$

L'homogénéité locale

$$HM_{d,\theta} = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} \frac{p(i, j, d, \theta)}{1 + |i - j|} \quad (2.89)$$

L'entropie

$$ET_{d,\theta} = - \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} p(i, j, d, \theta) \log(p(i, j, d, \theta)) \quad (2.90)$$

La probabilité maximale

$$PM_{d,\theta} = \max_{ij} \{p(i, j, d, \theta)\} \quad (2.91)$$

Dans ces formules,  $N_G$  désigne le nombre de niveaux de gris et  $\mu_x, \sigma_x, \mu_y, \sigma_y$  représentent respectivement les moyennes et écarts type des distributions marginales des éléments de la matrice de co-occurrence.

$$\mu_x = \sum_{i=1}^{N_G} i \sum_{j=1}^{N_G} p(i, j, d, \theta) \quad (2.92)$$

$$\mu_y = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} j p(i, j, d, \theta) \quad (2.93)$$

Cette méthode est très utilisée en analyse d'images. On en trouvera des applications dans [57] [158] [163] [197] [278] [285] [322].

Des études comparatives non exhaustives [73] ont démontré son efficacité et sa capacité de sélection et de caractérisation.

Le principal inconvénient de cette méthode réside dans le temps de calcul requis pour des images de grandes tailles et ayant un nombre important de niveaux de gris ainsi que dans la difficulté de choix du vecteur de déplacement.

### 2.2.3.2.3 Matrices de longueurs de plages

Dans [134], Galloway a introduit la méthode de matrice de longueurs de plages, où une plage est définie comme un ensemble consécutif de points ayant des niveaux de gris similaires, dans une direction donnée. Une plage est caractérisée par sa longueur et son orientation  $\theta$  (principalement 0, 45, 90 et 135 degrés). On définit alors les éléments de la matrice des longueurs de plages  $L_p$  de taille  $N_G \times L$  (où  $L$  la longueur maximale d'une plage), par:

$L_p(i, l, \theta)$  est égale au nombre de fois où l'image contient une plage de niveau de gris  $i$  de longueur  $l$  dans la direction  $\theta$ .

Comme c'est le cas pour les matrices de co-occurrence, on peut calculer les matrices de longueurs de plages pour différents angles.

De ces matrices Galloway propose d'extraire les paramètres suivants.

Importance des plages courtes

$$SER = \sum_{i=1}^{N_G} \sum_{l=1}^L [L_p(i, l, \theta) / l^2] \quad (2.94)$$

Importance de plages longues

$$LRE = \sum_{i=1}^{N_G} \sum_{l=1}^L [l^2 L_p(i, l, \theta)] \quad (2.95)$$

Distribution des niveaux de gris

$$GLD = \sum_{i=1}^{N_G} [\sum_{l=1}^L L_p(i, l, \theta)]^2 \quad (2.96)$$

Distribution des longueurs de plages

$$RLD = \sum_{l=1}^L [\sum_{i=1}^{N_G} L_p(i, l, \theta)]^2 \quad (2.97)$$

Pourcentage de plages

$$RLP = \sum_{i=1}^{N_G} \sum_{l=1}^L L_p(i, l, \theta) \quad (2.98)$$

Le paramètre *SRE* (respectivement *LRE*) est révélateur de la présence de texture fine (respectivement grossière). Il est d'autant plus élevé que la probabilité de rencontrer de plages courtes est élevée. Le paramètre *GLD* sera d'autant plus petit que l'histogramme de répartition des plages pour différents niveaux de gris sera plus uniforme. Inversement une répartition hétérogène induira une valeur élevée du paramètre. Le paramètre *RLD* mesure la non uniformité de la répartition des longueurs de plages indépendamment de leurs niveaux de gris. Enfin le paramètre *RLP* mesure le nombre de plages par unité de surface. Il est d'autant plus faible que les plages sont longues.

On trouvera dans [99] [130] des applications de cette méthode. Chu et al. introduisent deux autres paramètres spécifiques à des intensités de luminances particulières:

Importance des plages de niveaux de gris faibles

$$LGRE = \sum_{i=1}^{N_G} \sum_{l=1}^L [L_p(i, l, \theta)/i^2] \quad (2.99)$$

Importance des plages de niveaux de gris élevés

$$HGRE = \sum_{i=1}^{N_G} \sum_{l=1}^L [i^2 L_p(i, l, \theta)] \quad (2.100)$$

D'autres descripteurs ont été définis. On remarque, de part leur définition et au travers des applications qui s'y portent, qu'ils sont plus redondants que complémentaires. Il s'avère donc inutile de les utiliser dans leur totalité.

Notons toutefois que si cette méthode donne des matrices de taille plus faible que celles des co-occurrences, les caractéristiques extraites de la matrice de longueurs de plages varient énormément en fonction de  $l$ . Ceci dénote une certaine instabilité, et oblige à calculer plusieurs matrices, pour des distances différentes entre les niveaux de gris.

Une fois les paramètres extraites, différentes techniques de classification peuvent être employées (les plus proches voisins, nuées dynamiques,...).

Une évaluation des performances de classification sur un ensemble de textures [73], entre les méthodes basées sur les matrices de co-occurrence, les longueurs de plages, les différences de niveaux de gris et les descripteurs de Fourier, montre la supériorité des paramètres de la matrice de co-occurrence.

#### 2.2.3.2.4 Analyse de texture par filtrage

- **Filtrage spatial**

Le filtrage spatial est également très utilisé pour la caractérisation des textures. Dans [205], Laws propose l'utilisation d'un banc de filtres de tailles et de formes différentes pour produire une suite d'images qu'il appelle images d'"attributs de microtextures". Dans une seconde étape, il utilise des mesures comme la variance ou l'écart type pour calculer sur chacune des images produites lors des filtrages précédents ce qu'il désigne par les "attributs de macrostatistique". Ces derniers sont compressés en utilisant une analyse en composantes principales avant d'être envoyés au classifieur de texture.

On trouvera dans [84] une extension de la méthode de Laws s'appuyant non pas sur l'utilisation de masques fixes mais sur des masques dont les coefficients sont déterminés de manière adaptative.

Une formalisation générale de l'analyse de texture par filtrage spatial est présentée dans [339]. L'auteur utilise un formalisme vectoriel et présente l'analyse de texture par filtrage linéaire. Le choix judicieux du filtrage impose deux conditions. La première implique la préservation de l'énergie totale de la texture transformée (cette énergie est localisée dans les termes de la diagonale de la matrice de covariance spatiale de la texture). La seconde condition impose l'orthogonalité des valeurs de la base de transformation, ceci dans le but de garantir une décorrélation des composantes de l'image. Parmi les transformations répondant aux critères précédents on trouve celle de Karhunen-Loève.

Une extension de cette méthode consiste à l'utiliser conjointement avec des opérateurs non linéaires. Dans [337] l'auteur propose l'utilisation d'une combinaison de filtres spatiaux et d'opérateurs non linéaires pour caractériser la texture. Les paramètres locaux de texture sont obtenus par une succession de quatre opérations:

- transformation linéaire permettant la détection des structures locales;

- première transformation non linéaire de la forme  $f(x) = |x|a$ ;
- lissage itératif;
- seconde transformation homomorphique  $g(x) = \log |x|$ .

### • Filtrage fréquentiel

Un filtre de Gabor  $h$ , à deux dimensions, peut être représenté comme une fonction gaussienne modulée par une onde plane sinusoïdale [112]:

$$h(x, y) = \exp\left(-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right) \cos(2\pi(u_0x + v_0y) + \Phi) \quad (2.101)$$

où

$u_0, v_0$  et  $\Phi$  sont respectivement les fréquences et la phase de l'onde plane sinusoïdale.

$\sigma_x^2$  et  $\sigma_y^2$  caractérisent l'étendue spatiale du filtre.

Un filtre de Gabor d'orientation arbitraire  $\theta$  peut ensuite être obtenu en faisant subir une rotation au système d'axes  $(x, y)$ . En faisant varier les différents paramètres, plusieurs filtres de réponse impulsionnelle  $h_j$  sont obtenus. Pour une image  $f$ , le résultat  $f_j$  est déterminé à la sortie de chacun de ces filtres. Plusieurs caractéristiques peuvent ensuite être extraites de ces images comme, par exemple, la moyenne absolue de la déviation:

$$m_j = 1/n^2 \sum_{x,y}^{n-1} |f_j(x, y)| \quad (2.102)$$

Dunn et al. dans [112] utilisent ces filtres de manière différente: un seul filtre de Gabor (utilisé pour la nature de sa bande passante) sert à caractériser la texture. Ces travaux montrent que suivant les différences de texture, la forme des discontinuités peut prendre quatre types de signature différents. Des expériences ont montré que celles-ci constituaient un puissant outil pour segmenter des images.

Une autre approche proposée par Mallat [221] consiste à extraire des paramètres de texture en utilisant les décompositions en ondelettes. En effet, elles offrent la possibilité de mesurer des caractéristiques présentes à différentes échelles.

Une autre méthode proposée dans [335] consiste à utiliser plusieurs filtres exponentiels (en relation avec la perception visuelle humaine). Chacun des filtres correspond à un canal décomposant l'image en bandes de fréquences et d'orientation. A la sortie de chaque canal sont calculés un ensemble de quatre paramètres de texture mesurant l'énergie spectrale. Une approche similaire est présentée dans [24]. Elle est basée sur l'utilisation d'un banc de filtres de Gabor symétriques et pairs. Les attributs de la texture sont obtenus à partir des sorties des filtres par transformation non linéaire et d'un calcul d'une mesure d'énergie autour de

chaque point. La segmentation est réalisée par un algorithme de fusion en utilisant le critère d'erreur quadratique. Les reproches qui sont faits à ces méthodes résident particulièrement dans le choix empirique des filtres ainsi que dans la redondance d'information engendrée par la corrélation des images filtrées.

### 2.2.3.2.5 Paramètres géométriques

Une autre méthode pour caractériser des textures consiste à utiliser la dimension fractale [263]. En effet, certaines études ont mis en évidence la correspondance entre la dimension fractale et l'évaluation humaine de la rugosité. De plus, Pentland [262] a montré que si une surface naturelle est fractale, alors l'image d'intensité lumineuse de cet objet l'est aussi. La dimension fractale d'un objet peut donc être calculée en utilisant les niveaux de gris qui lui correspondent. Trois grandes méthodes permettant de calculer cette dimension fractale existent dans la littérature: la méthode des variations [28]; la méthode des boîtes [185]; et la méthode du spectre de puissance [262].

Dans le cas de l'approche morphologique, les formes analysées sont les formes des motifs constituant la texture. Cette approche a été proposée par Matheron [225] et Serra et Varchery [321]. Elle requiert la définition d'un élément structurant (ensemble de point ayant une forme spécifique telle qu'une ligne, un disque ou un carré) et la génération d'une image binaire, qui résulte de la translation de l'élément structurant au travers de l'image et l'érosion de celle-ci par l'élément structurant. La caractérisation de la texture peut être obtenue à partir de la nouvelle image binaire par accumulation des points ayant la valeur 1. Une application de cette approche est présentée dans [131].

Dans [89], Chen et al. proposent d'extraire 16 paramètres de texture à partir d'attributs géométriques. L'image, codée sur  $N_G$  niveaux de gris, est seuillée avec un seuil  $s$  variant entre 0 et  $N_G - 1$ .  $N_G$  images binaires sont ainsi obtenues. A partir de celles-ci, 16 paramètres géométriques relevant de la forme des régions 0 et 1 des images, sont extraits. Ils sont fondés sur des notions de compacité et d'irrégularité. Chen et al. montrent dans leur étude que ces paramètres sont aussi discriminants, sinon plus que d'autres classiquement utilisés dans la littérature.

### 2.2.3.3 Approche par modélisation de texture

Les approches par modélisation stochastique sont actuellement très utilisées en analyse d'images. L'utilisation de modèle en segmentation d'images ne se limite pas aux seules données image mais considère conjointement des connaissances a priori et des données contextuelles.

Nous abordons dans la première partie de cette section la modélisation linéaire de la texture. Nous présentons ensuite la modélisation par champ de Gauss-Markov.

### 2.2.3.3.1 Modélisation linéaire de la texture

Dans ce type de modélisation, la valeur du niveau de gris de chaque pixel est considérée comme une combinaison linéaire de celle des pixels voisins. C'est l'idée de base des modèles de prédiction linéaire 2D.

Une image  $I = \{I(i, j) / \text{tel que } 0 \leq i, j \leq M - 1\}$  est, dans ce cas, modélisée par:

$$I(i, j) = \sum_{(k, l) \in R} \alpha(k, l) I(i - k, j - l) + \sum_{(k', l') \in R} \beta(k', l') \epsilon(i - k', j - l') \quad (2.103)$$

$\epsilon(i, j)$  est un élément d'un champ aléatoire  $\epsilon$ ;  $R$  est la région de prédiction.  $M$  représente les tailles horizontale et verticale de l'image.

Selon le choix de la forme de  $R$ , on obtient trois régions canoniques: la région causale NSHP "Non Symetric Half Plane", la région semi-causale et la région non-causale (Fig.2.6).

FIG. 2.6 - Causalité dans les modèles 2D

Différentes modélisations linéaires peuvent être employées [61] [175] pour décrire un signal 2D. Dans les paragraphes suivants nous résumons le cas des modélisations AR, MA et ARMA.

#### • Modèle AR "Auto-Regressive"

Le principe du modèle autoregressif est d'utiliser la fonction d'autocorrélation comme mesure de la dépendance linéaire d'un point de texture par rapport à ses voisins [232].

On est dans le cas particulier de la modélisation linéaire où:

$$\begin{cases} \beta(0, 0) = 1 \\ \beta(k', l') = 0, \forall (k', l') \neq (0, 0) \end{cases}$$

c'est à dire

$$I(i, j) = \sum_{(k, l) \in R} \alpha(k, l) \cdot I(i - k, j - l) + \epsilon(i, j) = \hat{I}(i, j) + \epsilon(i, j) \quad (2.104)$$

$\hat{I}(i, j)$  correspond à l'estimée de  $I(i, j)$  et  $\epsilon$  représente l'erreur de prédiction (supposée centrée).

- **Modèle MA "Moving Average"**

Le modèle MA d'un signal 2D s'écrit sous la forme :

$$I(i, j) = \sum_{(k,l) \in R} \beta(k, l) \epsilon(i - k, j - l) \quad (2.105)$$

On est alors dans le cas particulier de la modélisation linéaire où  $\alpha(i, j) = 0, \forall (i, j) \neq (0, 0)$ .

Le modèle ARMA correspond au couplage des deux modèles précédents.

- **Estimation des paramètres du modèle AR**

Pour déterminer les coefficients  $\alpha(k, l)$  du modèle AR, il existe un grand nombre d'algorithmes, parmi lesquels on peut citer l'autocorrélation (algorithme de Yule-Walker), la covariance, la covariance modifiée, la corrélation partielle (algorithme de Burg), le maximum de vraisemblance [175]. Les performances de ces différents algorithmes sont tout à fait comparables lorsqu'il s'agit d'un ensemble de données assez important.

Dans le cas du critère des moindres carrés, les coefficients sont calculés en minimisant l'espérance du carré de l'erreur de prédiction.

L'ensemble d'équations linéaires à résoudre, dans le cas 2D, est de la forme:

$$\gamma(m, n) - \sum_{(k,l) \in W_R} \alpha(k, l) \cdot \gamma(m - k, n - l) \text{ pour } (0 \leq m, n \leq M - 1) \quad (2.106)$$

où

$$\gamma(m, n) \equiv E[I(i, j) \cdot I(i - m, j - n)] \quad (2.107)$$

La résolution de ce système permet alors d'obtenir les  $\alpha(k, l)$ .

Dans le cas où  $\gamma(m, n)$  est inconnue, son estimation peut être faite à l'aide de l'une des deux méthodes suivantes:

- **Méthode d'auto-corrélation**

Les données sont supposées nulles hors un ensemble ou une fenêtre défini de points noté  $W_S$ . On obtient alors:

$$\gamma(m, n) = \sum_{(i,j) \in W_R, (i-m,j-n) \in W_S} I(i, j) \cdot I(i - m, j - n) \quad (2.108)$$

### – Méthode de corrélation

Les données ne doivent pas se trouver en dehors d'une fenêtre prédéfinie  $W_0$ . Dans ce cas, on a :

$$\gamma[(0, 0), (-m, -n)] - \sum_{(k,l) \in W_R} \alpha(k, l) \cdot \gamma(-k, -l) = 0 \quad (2.109)$$

et

$$\gamma[(-k, -l), (-m, -n)] = \sum_{(k,l) \in W_0} I(i - k, j - l) \cdot I(i - m, j - n) \quad (2.110)$$

En pratique, il existe principalement deux classes d'algorithmes: direct et itératif. Parmi les méthodes directes, l'algorithme de Levinson-Durbin est le plus efficace. Pour les méthodes itératives, on peut citer la méthode du gradient, la méthode de Newton, la méthode stochastique [175].

Les coefficients  $\alpha(k, l)$  du modèle AR convergent vers des valeurs stables qui ne dépendent plus que de la position du point traité. Ils sont utilisés pour caractériser la texture.

#### 2.2.3.3.2 Modélisation par champs de Markov

Une classe des modèles se prêtant bien pour la prise en compte des interactions locales aux niveaux des points est la classe des champs de Markov. Son application concerne la restauration d'images [28] [56] [138], la segmentation bayésienne [65] [106], la synthèse de textures [79], la reconstruction d'images [223], etc...

La plupart de ces tâches peuvent se formaliser dans un cadre général appelé l'étiquetage d'images où à chaque pixel, on veut associer une étiquette (label) appartenant à un ensemble fini. Le problème est de choisir une étiquette optimale pour un pixel.

Un des intérêts majeurs du champ de Markov est sa capacité à fournir un modèle mathématique simple. Les caractéristiques à un site donné dépendent uniquement des interactions locales. D'autre part, le succès de son utilisation est due en grande partie au théorème de Hammersley-Clifford [138] qui permet de simplifier l'écriture des probabilités conditionnelles et faciliter la détermination des paramètres qui spécifient le modèle. Les interactions au niveau du point sont prises en compte par le biais d'énergies potentielles.

La segmentation, associée aux champs de Markov et à la théorie bayésienne, se ramène à un problème d'optimisation. Le but est alors de trouver le minimum global d'une fonction représentant l'erreur commise (appelée également énergie). Celle-ci décrit les interactions entre les différentes variables modélisant les caractéristiques images d'un problème donné. Deux types de variables sont considérés: les variables correspondant à une représentation des données initiales (images), et celles relatives aux étiquettes (régularisation).

Dans la première partie de ce paragraphe, nous présentons des généralités sur les champs de Markov. Leur application au traitement d'images et plus particulièrement à la segmentation est introduite dans un second temps. Le problème de la minimisation globale de la fonction d'énergie est ensuite abordé par l'introduction de plusieurs algorithmes de relaxation. Un sous-paragraphe est consacré à l'estimation des paramètres dans le cadre de la segmentation non supervisée des textures. Enfin, quelques stratégies de relaxation sont présentées.

- **Généralités**

- **Système de voisinage**

Soit  $S = \{\mathbf{s} = (i, j); 1 \leq i \leq N; 1 \leq j \leq N\}$  un ensemble de sites. Soit  $G = \{G_s | \mathbf{s} \in S\}$  un système de voisinage défini sur la grille  $S$ , tel que:

- $\mathbf{s} \notin G_s$
- $\mathbf{s} \in G_r \equiv r \in G_s$

Les éléments de  $G_s$  définissent les sites voisins de  $\mathbf{s}$ . Un sous-ensemble  $c$  de  $S$  est appelé une clique relative au système de voisinage  $G$  si  $c$  représente un singleton ou si tous les sites distincts de  $c$  sont voisins deux à deux. Une clique est dite d'ordre  $n$  si elle contient  $n$  éléments. Dans la figure 2.7 est représenté l'ensemble des cliques  $c$  associées à un système de voisinage du premier et du second ordre. L'ensemble des cliques de  $S$  sera noté  $C$ . Pour des raisons de temps de calcul, les cliques d'ordre deux sont les plus couramment utilisées.

FIG. 2.7 - *Système de voisinage et cliques associées*

- **Champ de Markov et distribution de Gibbs**

Soit  $E = \{e_s, \mathbf{s} \in S\}$  un ensemble de variables aléatoires indexées par  $\mathbf{s}$  et à valeur dans un ensemble fini  $\Lambda = \{0, 1, \dots, L - 1\}$ .  $E$  représente le champ des étiquettes ou des labels à estimer.

Soit  $(E = \omega)$  un événement tel que  $\{e_1 = \omega_1, \dots, e_N = \omega_N\}$  avec  $\omega = \{\omega_1, \dots, \omega_i, \dots, \omega_N\}$  et  $\omega_i \in \Lambda$ .

Soit  $\Omega$ , l'ensemble de toutes les configurations possibles de  $\omega$ ,  $\Omega = \{\omega = (\omega_1, \dots, \omega_i, \dots, \omega_N); \omega_i \in \Lambda\}$ .

Le vecteur aléatoire  $E$  est un champ de Markov, relativement au système de voisinage  $G$  si:

- $\forall \omega \in \Omega; P(E = \omega) > 0$
- $\forall \mathbf{s} \in S,$

$$P(e_s = \omega_s | e_r = \omega_r; r \in S - \{\mathbf{s}\}) = P(e_s = \omega_s | e_r = \omega_r; r \in G_s) \quad (2.111)$$

Cette considération fait que chaque élément aura uniquement une interaction avec ses voisins. D'autre part, le théorème de Hammersley-Clifford [138], relatif à la distribution de Gibbs, montre qu'un champ aléatoire  $E$  est un champ markovien associé au système de voisinage  $G$ , si et seulement si sa distribution de probabilité  $P(E = \omega)$  est une distribution de Gibbs définie par:

$$\forall \omega \in \Omega; P(E = \omega) = \frac{1}{Z} \exp\{-U(\omega)\} \quad (2.112)$$

où  $Z$  est une constante de normalisation appelée fonction de partition de Gibbs.

$$Z = \sum_{\omega \in \Omega} \exp\{-U(\omega)\} \quad (2.113)$$

La fonction d'énergie  $U$  peut être définie de manière locale par:

$$U(E = \omega) = \sum_{c \in C} V_c(\omega) \quad (2.114)$$

où  $C$  est l'ensemble des cliques relatives au système de voisinage  $G$ ;

$V_c$  est une fonction définie sur  $\Omega$ , dépendant uniquement des éléments  $\omega_s$  tels que  $\mathbf{s} \in C$ .

Cette fonction est appelée fonction de potentiel. Grâce aux fonctions de potentiel, nous pouvons définir de manière simple les propriétés locales du champ  $E$ .

### • Application à la segmentation d'image

Nous recherchons maintenant un modèle pour la segmentation d'images qui consiste à affecter une étiquette à chaque site (le pixel) de la scène. Les données sont un ensemble de sites  $S$  composant l'image, un système de voisinage  $G$  et un ensemble d'observations (données image)  $O = \{o_s; \mathbf{s} \in S\}$ . Chacun des sites  $\mathbf{s}$  doit être assigné à une étiquette à valeur dans un ensemble  $\Lambda$ . Nous supposons que le champ des étiquettes  $E$  est un champ aléatoire markovien. Le problème est alors de trouver la segmentation optimale  $\omega^*$ .

Afin de trouver une segmentation optimale  $\omega^*$ , plusieurs estimateurs sont décrits dans la littérature, chacun d'entre eux correspondant à une fonction de coût différente. Ils consistent à chercher un estimateur  $\hat{\omega}$  qui maximise la probabilité  $P(E = \omega / O = o)$ .

### – Segmentation supervisée

L'estimateur le plus souvent retenu est celui du Maximum A Posteriori MAP. Celui-ci consiste à chercher un estimateur  $\hat{\omega}$  qui maximise la probabilité  $P(E = \omega/O = o)$  et donc, qui minimise l'énergie  $U(E = \omega/O = o)$ .

Un autre estimateur, le “Maximum Posterior Marginal” MPM est généralement employé. Cet estimateur considère que le nombre de sites  $\mathbf{s}$  mal étiquetés est une meilleure mesure de la fonction coût. L'estimateur est donc:  $\hat{\omega} = (\hat{\omega}_1, \hat{\omega}_2, \dots, \hat{\omega}_N)$  où chaque  $\hat{\omega}_i$  maximise de manière locale  $P_s(e_j/O = o) \quad 0 \leq j < L$ .  $P_s(e_j/O = o)$  représente la probabilité de la marginale a posteriori d'avoir l'étiquette  $e_j$  au site  $\mathbf{s}$ . Cette probabilité est conditionnée par l'observation  $O$  et non pas par l'observation du voisinage de  $\mathbf{s}$  dans  $S$ , ce qui entraîne des calculs assez lourds.

Notons que d'autres estimateurs tels que le “Middle Field” MF et le “Threshold Posterior Mean” TPM peuvent être utilisés.

La minimisation de  $U(E = \omega/O = o)$  est un problème d'optimisation difficile en raison de la dimension souvent élevée de l'ensemble de toutes les configurations possibles de  $E$  et de la présence possible de nombreux minima locaux.

Plusieurs méthodes d'optimisation sont alors envisageables:

Des méthodes stochastiques comme décrites dans [138] [343] convergent asymptotiquement vers le minimum global, mais nécessitent un temps de calcul élevé.

Parmi ces méthodes on trouvera l'algorithme de Metropolis [240] ou l'échantillonneur de Gibbs [138].

D'autres méthodes déterministes convergent vers des minima locaux. Elles sont sous optimales mais plus rapides que les méthodes stochastiques.

A titre d'exemples, des méthodes déterministes on trouve l'algorithme GNC “Graduated Non Convexity” [53] [42] [291], ICM “Iterated Conditional Mode”, HCF “Highest Confidence First” [28] ou encore le MFA “Mean Field Annealing” [139].

### – Segmentation non supervisée

Les approches par modélisation statistique conduisent à de bons résultats lorsque les paramètres des modèles statistiques et le nombre de régions à extraire sont connus a priori.

Dans le cas de la segmentation non supervisée, les paramètres formant un vecteur  $\theta$ , définissant la loi  $P(E/O)$  ne sont pas connus a priori. Le problème se ramène alors à la

recherche d'estimateurs de  $\omega$  et de  $\theta$  tels que:

$$(\hat{\omega}, \hat{\theta}) = \mathit{Arg} \max_{\omega, \theta} P(E = \omega/O, \theta) \quad (2.115)$$

Cette maximisation est très difficile à obtenir et n'admet pas de solution directe. Une solution possible à adopter pourra alors être [207]:

$$\hat{\omega} = \mathit{Arg} \max_{\omega \in \Omega} P(E = \omega/O, \hat{\theta}) \quad (2.116)$$

$$\hat{\theta} = \mathit{Arg} \max_{\theta} P(E = \hat{\omega}/O, \theta) \quad (2.117)$$

Pour résoudre ce problème, Dempster et al. [115] ont développé l'algorithme EM "Expectation Maximisation". Il permet, à partir d'une valeur initiale  $\theta^0$  des paramètres, de définir une suite  $\theta^q$  convergeant vers le paramètre pour lequel la vraisemblance par rapport aux données est maximale. D'autres algorithmes permettent d'obtenir de manière statistique les paramètres du modèle comme, par exemple, l'algorithme ICE (Iterative Conditional Estimation) proposé par Pieczynski [268] et utilisé par Kato et al. [203] pour réaliser de la classification non supervisée. Il existe également l'algorithme SEM "Stochastic Expectation Maximization" défini par Celeux et al. [68].

Une autre façon d'estimer ces paramètres a été proposée par Manjunath et Chellappa [227] dans le contexte d'une segmentation non supervisée d'images texturées. L'image est divisée selon une grille et les coefficients de texture sont déterminés pour chaque fenêtre ainsi créée. Un processus de classification dépendant d'un seuil est alors engagé afin de trouver les paramètres de texture correspondant aux différentes classes. Une fois le nombre de classes de texture et leurs paramètres connus, un processus de relaxation markovienne classique est utilisé. Le problème de cet algorithme est que le processus de classification dépend lui-même d'un seuil. D'autre part, une mauvaise classification entraînera de mauvais paramètres, et par la suite, une mauvaise segmentation de l'image.

D'autres travaux sur l'estimation des paramètres se trouvent dans [104] [140] [190] [207] [364]

Par ailleurs, il existe de nombreuses manières d'utiliser les champs de Markov en segmentation d'images, qui dépendent de la façon dont le champ d'observation est considéré. Parmi celles-ci, on peut citer l'approche de base [106] [190] [207] et [359], l'approche fondée sur les graphes d'adjacence [4] [202] et [243], l'approche multi-échelles [7] [50] et [266] et l'approche hiérarchique [203].

#### 2.2.3.4 Conclusion

Sans être exhaustif, nous avons examiné la plupart des méthodes de segmentation statistiques en les classant en deux catégories: celles basées sur l'extraction de paramètres

puis sur la classification et celles basées sur l'adaptation d'un modèle de texture à l'image.

La première catégorie de méthodes (notamment les co-occurrences, les longueurs de plages et les dépendances spatiales) s'applique correctement à une grande variété d'images et permet une bonne caractérisation à la fois des micro et macro-textures. Les paramètres de ces méthodes (vecteur de déplacement, seuils, tailles de fenêtres) sont toutefois, souvent difficiles à déterminer et exigent dans le cas des images de grande taille et avec un nombre important de niveaux de gris, un temps de calcul très important.

La seconde famille de méthodes sont puissants pour analyser certaines textures. Le modèle markovien est utilisé avec beaucoup de succès dans de nombreux domaines d'analyse d'images. Il permet une modélisation simple et pertinente des contextes et interactions locales de pixels, mais exige un nombre considérable d'itérations pour converger vers des solutions stables ce qui augmente le temps de calcul. Sa performance s'avère néanmoins limitée dans le cas des textures à caractère régulier.

## 2.3 Approches coopératives

### 2.3.1 Introduction

Plusieurs approches coopératives de segmentation, fondées sur la prise en compte de différentes formes de représentation de l'image, mais aussi sur l'application de différentes formes de traitement peuvent être rencontrées dans la littérature. Dans le cas général, nous supposons disposer de  $N_r$  représentations de la même image obtenues soit par des méthodes différentes, soit par la même méthode, mais avec des paramètres différents. Les représentations qui restent le plus souvent manipulées dans la majorité des approches coopératives sont celles des contours et des régions. Ces deux types de primitives se fondent respectivement sur des notions de dissimilarité et de similarité des propriétés de l'image. Elles sont donc théoriquement duales et ont des caractéristiques complémentaires. Toutefois, en pratique, les représentations obtenues à l'aide de chacune de ces méthodes prise séparément ne vérifient que très rarement la complémentarité précédente. Chacune des méthodes apporte une contribution partielle au problème, et pour obtenir de meilleurs résultats il est nécessaire de mettre en oeuvre un procédé permettant la coopération de l'information "contour" et de l'information "région".

Trois schémas de traitements sont souvent empruntés pour mettre en évidence l'aspect coopératif de la segmentation : le **schéma séquentiel**, le **schéma parallèle** et le **schéma hybride**. Le premier concerne les méthodes où l'extraction des contours est préalablement effectuée pour contrôler le processus de croissance de régions. Le second est celui des méthodes dans lesquelles il y a une extraction conjointe et simultanée des contours et des régions. Enfin, le schéma hybride combine les deux précédents.

Quelques exemples illustrant différentes démarches de coopération entre plusieurs types de primitives sont résumés dans les paragraphes suivants.

## 2.3.2 Approches séquentielles

Dans les systèmes de segmentation opérant par coopération séquentielle entre détection des régions et détection des contours, ces derniers sont en général utilisés comme des contraintes, ou des références, permettant de guider et de contrôler les processus d'extraction des régions.

### 2.3.2.1 Contrôle de la fusion des régions par les contours

Wrobel et Monga [356] proposent une technique de segmentation “régions” guidée par des points de contours obtenus par extraction préalable. Leur algorithme fonctionne en trois étapes. La première consiste à réaliser une segmentation à l'aide d'une procédure de type fusion de régions [258] sur critère *Min - Max*. Le résultat de cette segmentation est ensuite représenté sous forme de graphe d'adjacence lors de la seconde phase. Chaque noeud de graphe est associé à une région et les arcs représentent les relations d'adjacence entre régions. La dernière étape consiste à regrouper les couples de régions voisines sur la base d'un critère dépendant de la variance de la région totale et du pourcentage de points contours sur la frontière commune aux deux régions à fusionner.

Dans le but de la mise en correspondance de régions d'images stéréoscopiques, une approche similaire est proposée dans [340]. On trouve une autre approche dans [196], où une fusion brutale est proposée dans laquelle la segmentation en contour est superposée à l'image des régions. Ainsi, les contours qui sont prioritaires par rapport aux régions servent à diviser les régions sous segmentées.

### 2.3.2.2 Coopération contours-modèle surfacique

Dans [136], Gombotto propose une méthode d'extraction de régions à partir d'un germe initial placé par un utilisateur dans l'image. L'agrégation des pixels au germe se fait d'une façon itérative. A chaque itération, les pixels adjacents à la région en cours de formation sont d'abord groupés en segments homogènes (un segment représente une ligne d'épaisseur d'un pixel). Un segment est annexé à la région s'il est statistiquement proche du modèle de la région. La région est modélisée par une surface plane bruitée :

$$I(i, j) = Ai + Bj + C + n_{i,j} \quad (2.118)$$

où  $I(i, j)$  représente le niveau de gris du pixel  $(i, j)$ ,  $A$ ,  $B$  et  $C$  sont des constantes, et  $n_{i,j}$  est un bruit blanc centré. L'arrêt de la croissance est lié à une information de type contour qui est le gradient moyen calculé à la frontière de la région en formation. En fait la croissance est effectuée sur un grand nombre d'itérations et la frontière de la région recherchée est celle qui donne le maximum du gradient. Cette procédure est similaire au modèle du contour actifs

[201] dans le sens où l'on cherche la frontière qui maximise une fonction de gradient. Par contre aucune information sur la courbure du contour n'est utilisée pour assurer un contour lisse.

### 2.3.2.3 Coopération région-région et région-contour

Récemment, Kara-Falah et al. [184] ont proposé une technique originale d'extraction des germes de croissance. Le principe peut être résumé comme suit : après la réalisation de plusieurs segmentations, une segmentation (consensus) est obtenue. Les noyaux des régions de cette segmentation sont ensuite considérés comme candidats pour former les germes. Ceux-ci sont alors soumis à un test de stationnarité et seuls ceux correspondant à une région homogène sont conservés. Cette première étape est nommée coopération région-région. Le processus de croissance débute ensuite : les pixels voisins des germes sont agglomérés de façon itérative s'ils vérifient le critère d'homogénéité défini. Ici, la croissance repose à la fois sur des propriétés globales (moyenne des niveaux de gris,...) et sur des informations provenant des contours. Ceci permet d'éviter d'éventuels conflits aux frontières des régions. Cette deuxième étape plus classique constitue la coopération région-contour.

Dans [365], Xiaohan et al. utilisent un principe très proche de la seconde étape de la coopération précédente. Ils intègrent directement l'information sur les gradients au sein même du mécanisme de choix du pixel à agréger dans une croissance de région incrémentale.

### 2.3.2.4 Coopération division-fusion et contours actifs

Pavlidis et Liow [270] proposent une approche quelque peu différente combinant un algorithme de division-fusion avec un processus de détection de contours. Ils utilisent l'avantage de la technique de croissance de régions afin d'obtenir des contours fermés, et corrigent ensuite leur position par une méthode de déformation de contours. La méthode procède en 3 phases :

- Division-fusion par *quadtree* avec un seuillage qui produit une sur-segmentation.
- Elimination des frontières; chaque frontière entre deux régions est conservée ou enlevée en fonction de son contraste, ses changements de direction et sa longueur.
- Modification des contours; les frontières conservées lors de la phase précédente sont considérées comme des contours significatifs. Leurs positions sont corrigées à l'aide d'une approche par contour actif [201].

### 2.3.2.5 Coopération squelettes-contour-région

Dans le but du codage d'images, une autre technique faisant coopérer les contours et les régions est proposée par Benois et Barba dans [23]. L'aspect coopératif de cette méthode

est dans le choix de points particuliers de squelettes, construits à partir des points contours, comme centres de croissance de régions. L'algorithme développé se déroule en trois étapes.

- Détection de contours.
- Squelettisation des régions définies par la détection de contours, et sélection à partir du squelette d'un certain nombre de points qui serviront de noyaux pour le processus de croissance de régions.
- Segmentation avec contrôle des points contours.

### 2.3.2.6 Coopération contour-région fondée sur le diagramme de Voronoï

Dans [228], une autre méthode est développée par Melkemi et Chassery. La coopération région-contours est fondée sur une représentation par diagramme de Voronoï. Les auteurs proposent une approche de partition de type division-fusion pour découper l'image en polygones de Voronoï homogènes. L'aspect de coopération est issu du fait d'assimiler des segments de contours extraits au préalable à des arêtes de polygones et de les utiliser pour contrôler la localisation des frontières des régions de Voronoï.

## 2.3.3 Approches parallèles

### 2.3.3.1 Intégration de l'extraction de région et de contour

Dans le but d'une application de reconstruction 3D, Bonnin et Zavidovique [54] proposent une méthode coopérative de segmentation d'images. Le processus de segmentation est composé de plusieurs étapes.

Les points de contours sont extraits les premiers. Leur position dans l'image est précise, aussi, s'ils ne sont pas détruits dans la suite de l'algorithme, leur position restera inchangée. Une première segmentation en régions est ensuite effectuée. Elle a pour rôle de produire des régions homogènes ne comportant aucun point de contour. Le module est composé d'une procédure de découpage itératif suivie d'une procédure de fusion. Les frontières des régions sont ensuite obtenues par fusion de pixels adjacents à l'intérieur des régions.

La seconde étape est guidée par les contours. Les frontières ainsi extraites sont suffisamment précises sur le plan de la localisation pour être confrontées aux points de contours.

La troisième étape est celle de la confrontation des résultats. Elle consiste à comparer les points de frontières de régions, de manière à éliminer les points de contours en trop, c'est-à-dire ceux qui ne sont pas situés à proximité d'une frontière de région. La fusion des informations duales consiste ici à confirmer l'information contour par l'information région.

Lors d'une quatrième étape, un module de chaînage des points de contours et prolongation à courte distance, permet de connecter les contours ouverts situés sur les frontières de régions.

La cinquième étape est celle de prolongement à longue distance et fusion finale de régions. Cette étape consiste à examiner la frontière commune de chaque paire de régions adjacentes puis à prendre l'une des deux décisions suivantes :

- fusion des régions adjacentes et effacement du contour sur la frontière commune.

- non fusion et prolongation du contour sur toute la frontière commune.

Cette étape garantit une compatibilité entre les deux segmentations produites. Pour terminer, un module de mise en forme des données remplit la structure de données des régions, avec les résultats compatibles des deux segmentations.

### 2.3.3.2 Extraction des régions et des contours par matrice de co-occurrence

Haddon et Boyce [158] effectuent une segmentation de l'image en régions et en contours par un partitionnement statistique de la matrice de co-occurrence. Ils exploitent la répartition spatiale des éléments de cette matrice et l'utilisent comme espace transformé de l'espace de l'image. Les auteurs définissent deux types de transitions de niveau de gris, d'une part les faibles transitions, celles situées à proximité de la diagonale principale de la matrice de co-occurrence, et qui sont caractéristiques des régions et, d'autre part les fortes transitions, celles situées hors diagonale et qui sont caractéristiques des points contours.

Une première étape de segmentation est effectuée selon le principe suivant. Un point de l'image est affecté à une région ou considéré comme point de contour en fonction de la position des transitions du niveau de gris avec son voisinage dans l'espace défini par la matrice de co-occurrence. La pertinence de la classification issue de l'étape précédente est ensuite vérifiée dans une seconde phase, sur la base d'un critère local utilisant une mesure d'entropie.

## 2.3.4 Approches hybrides

De nombreuses méthodes combinent les deux aspects de "détection-contour" et de "détection-région" en effectuant une fusion des résultats fournis par les deux approches.

### 2.3.4.1 Fusion de plusieurs segmentations

Une autre approche de coopération est proposé par Chabbi et Masini [55]. Elle opère par fusion des informations extraites à l'aide des détecteurs régions et contours. L'objectif de la méthode est d'extraire des facettes 2D, c'est-à-dire des chaînes fermées de segments de contours, projection dans l'image de facettes 3D d'objets polyédriques de la scène. Pour ce faire, les informations issues de la détection de contours et de régions sont combinées de la manière suivante :

- les segments sont, dans un premier temps, extraits de l'image, puis regroupés en chaînes connexes à partir de chaque paire de segments adjacents, sans aucun contrôle de l'information région. Pour chaque segment, l'information relative à la transition entre les deux régions qui le bordent est relevée.
- pour chaque chaîne connexe, plusieurs segmentations en régions sont envisagées, chacune ayant pour paramètre de contrôle de fusion (seuil) l'information de la transition

locale autour d'un des segments de la chaîne. Ces paramètres sont classés par ordre décroissant.

- une chaîne connexe est validée si elle n'englobe qu'une seule région, et si aucun de ses segments ne traverse une région. Ainsi, l'information de région permet soit de découper une chaîne connexe, soit de la fermer, de manière à obtenir un polygone fermé.

La deuxième phase de cette méthode contrôle l'extraction de régions par les contours, alors que la fusion intervient en troisième phase et permet d'extraire des polygones fermés dont l'intérieur est constitué d'une seule région.

La méthode de fusion proposée par Kara-Falah et Bolon [183] s'appuie, elle, sur une représentation en régions pour chaque segmentation, et cherche par un mécanisme d'association entre régions à mettre en évidence des ensembles de régions associées où chaque ensemble correspond à une région éventuelle dans l'image. Ceci permet de construire une segmentation floue où chaque région de l'image est définie comme une région floue. Cette représentation floue contient toute l'information fournie par les segmentations de départ. Le passage d'une segmentation floue à une segmentation nette permet de confirmer la partie redondante de l'information et de réduire l'importance de informations contradictoires. Une segmentation consensus est ainsi trouvée entre les différentes segmentations.

La construction de régions floues est effectuée de la manière suivante :

Soient  $R_1, R_2, \dots, R_n$ ,  $n$  sous-régions représentant une région  $R$  de l'image. On définit le noyau et le support de  $R$  par :

$$C_n(R) = \bigcap R_i \quad (2.119)$$

$$S_n(R) = \bigcup R_i \quad (2.120)$$

$C_n(R)$  représente l'ensemble de pixels qui appartiennent certainement à  $R$  et  $S_n(R)$  l'ensemble de pixels qui ont une possibilité d'appartenir à  $R$ .

Un certain degré de confiance peut être accordé pour chaque résultat de segmentation, ce qui revient à donner un poids  $m(R_i)$  pour chaque représentation  $R_i$  de façon à obtenir :

$$\sum_i m(R_i) = 1 \quad (2.121)$$

Une région  $R$  peut ainsi être définie comme une région floue avec une fonction d'appartenance  $\mu_R$  par :

$$\forall x \in I, \quad \mu_R(x) = \sum_{j/x \in R_j} m(R_j) \quad (2.122)$$

où  $I$  est l'ensemble des pixels de l'image.

Les auteurs utilisent les régions ainsi formées comme noyaux de la procédure de croissance régions. Le passage de la segmentation floue à une segmentation nette est appelé défuzzification. Ceci est opéré à l'aide de l'algorithme de relaxation proposé dans [371].

### 2.3.4.2 Coopération contour-région utilisant un système expert

Dans [247], Nazif et Levine proposent une coopération d'opérateurs de segmentation de bas niveau suite à la réalisation d'un système expert, dont les règles opèrent sur des connaissances de bas niveau, hors de toute sémantique liée à l'image. Ils réalisent tout d'abord une segmentation en régions. Des caractéristiques sont extraites pour chacune des régions ainsi obtenues telles que la moyenne et la variance des niveaux de gris, la taille, le périmètre, etc.

A partir des résultats d'une extraction de contours, on détermine un certain nombre de caractéristiques comme: la moyenne et la variance des points de contours, la moyenne et la variance le long d'une ligne, la direction moyenne d'un contour, etc. Une description spatiale de la scène est alors obtenue à partir des résultats précédents: région à droite ou à gauche d'une ligne,... Cette structure de données est examinée et modifiée avec un ensemble de règles [214] (quelques centaines).

Cette classe d'algorithmes donne de bons résultats, en particulier parce qu'elle repose sur la comparaison de deux segmentations déjà bien connues. Elle pallie à la fois les problèmes des algorithmes de fermeture de contours (régions avec des contours fortement contrastés, mais peu homogènes) et ceux des algorithmes de regroupement itératif (régions uniformes avec un faible contraste à leur contour) en fournissant des segmentations avec des régions uniformes, aux contours contrastés. Le gros problème de cette approche réside dans le choix des règles à appliquer. Celles-ci reposent souvent sur des notions qualitatives, par exemple, si la taille de la région est grande alors ... Elles nécessitent donc l'introduction de seuils dépendant des images. Un inconvénient des méthodes mixtes est qu'elles engendrent de forts coûts algorithmiques: même si les deux segmentations peuvent être réalisées en parallèle, le processus de haut niveau, fondé sur un ensemble de règles, est souvent long à réaliser.

### 2.3.4.3 Coopération contour-région par approche multi-agents

Dans cette approche chaque agent représente un petit système expert autonome lié aux autres agents. A titre d'exemple, Baujard et Garbay [35] proposent un système de vision pour l'identification de cellules biologiques, basé sur la coopération entre le détecteur de contours de Deriche [108] et la technique de classification de Fisher [125] appliquée aux régions, précédée d'un filtrage de l'image. Un agent est dédié à chacun des processus. Après détection des régions et des contours, les résultats sont analysés et comparés dans un agent spécifique. Lorsqu'il y a compatibilité entre la frontière des régions et les contours, le résultat est validé et envoyé à un agent qui tente de classer la région à partir de la forme de celle-ci. Si la compatibilité n'est pas vérifiée, les régions sont découpées ou fusionnées. Cette approche est particulièrement intéressante car elle pose clairement le problème de la gestion de la coopération. Elle est généralement jugée plus sur la manière d'aborder le problème que sur la qualité des résultats qu'elle engendre. Il y a de grandes difficultés à faire coopérer deux méthodes qui ont peu de points communs et qui ne sont pas spécialement préparées

pour cette coopération. Si le cumul des informations est important, la complémentarité des informations doit également être pris en compte pour que l'efficacité soit réelle. Ainsi, à partir d'un détecteur de contours, il est difficile de trouver un processus de détection des régions qui soit complémentaire. La distribution des rôles dans un univers multi-agents n'est donc pas remise en cause, mais elle n'est pas très efficace si les processus coopérants sont trop indépendants. Cette remarque reste valable pour les coopérations non fondées sur l'approche multi-agents. Il est difficile de cerner les avantages de l'approche multi-agents pour la coopération région-contours. Si chaque agent est distribué sur un processeur différent, il y a bien sûr un gain de temps.

#### 2.3.4.4 Extraction d'une segmentation consensus

Un travail récent sur la fusion des résultats de segmentation est celui de Chu et Aggarwal [58]. Cette technique part de plusieurs segmentations de l'image fournies en entrée (sous forme de segmentation contours). Elle calcule une segmentation consensuelle entre les segmentations données, sans avoir recours ni à une information donnée par un processus de haut niveau, ni à l'image originale. Par contre, une pondération peut être donnée à chaque segmentation fournie à l'entrée si l'utilisateur possède une information sur le degré de confiance accordé à chaque segmentation. L'algorithme est composé de trois étapes :

**Etape 1. Estimation initiale des points contours :** l'objectif de cette phase est d'estimer les vraies positions des points de contour à partir de plusieurs réalisations données par les segmentations initiales (qui sont converties en contours). L'estimation choisie est celle qui minimise au sens du maximum de vraisemblance, une fonction de coût de l'erreur d'estimation entre les échantillons en entrée et la sortie :

$$\sum_{T_i \in C} C_e(T_i, P_i) \quad (2.123)$$

où  $C$  désigne l'ensemble des points de contour en entrée.  $T_i$  désigne un point contour en entrée, et  $P_i$  désigne le point estimé correspondant.  $C_e(T_i, P_i)$  représente l'erreur d'estimation exprimée en terme de distance entre les deux points  $T_i$  et  $P_i$ . L'estimation est calculée localement. Pour chaque pixel de l'image  $x$ , les points en entrée sont extraits dans un cercle de rayon  $\rho$  centré sur le point examiné. Plusieurs points de contours provenant d'une même segmentation en entrée peuvent se trouver dans le cercle et donc participent à l'estimation. Pour avoir plus de précision dans l'estimation d'un point de contour  $P_i$  pour un cercle donné, les auteurs proposent de rejeter les points en entrée qui se trouvent à une distance supérieure à  $\rho$  du point  $P_i$  et de recommencer l'estimation. Ce processus est itéré tant qu'il y a des points rejetés. Une fois l'estimation point par point terminée, les points de contours en sortie peuvent ne pas former un contour fermé. La solution proposée consiste à connecter les deux points résultants de l'estimation effectuée sur deux cercles adjacents si leurs centres correspondent à des points de contours en entrée (si deux points de contours sont connexes en entrée, il faut qu'ils donnent des points connexes en sortie).

**Etape 2. Lissage du contour produit :** elle consiste à lisser le contour donné par la première phase de façon à minimiser une fonction de courbure tout en gardant une bonne

estimation. Le problème à résoudre est le suivant :

Soit  $\{P_j\}$  l'ensemble des points de contours, trouver un ensemble de points de contours  $\{Q_j\}$  qui minimise une fonction de courbure  $\sum_{j=1}^m C_s(Q_j)$  sous les contraintes suivantes :

- (a)  $d_e(P_j, Q_j) \leq \rho$  pour tout  $j$ ,
- (b)  $d_e^2(Q_i, Q_j) \leq c$  si  $d_e^2(P_i, P_j) \leq c$ .

$d_e(\cdot)$  désigne la distance euclidienne;  $c$  est une constante égale à 2, elle est utilisée pour représenter la 8-connexité. Les auteurs suggèrent une méthode itérative pour satisfaire les contraintes exprimées ci-dessous. Ils considèrent les points de contours comme des noeuds. Les noeuds adjacents sont reliés par des "ressorts". Les forces exercées par les ressorts expriment les contraintes de courbure et tendent à placer les noeuds dans une position qui minimise une fonction d'énergie. Les contraintes introduites en (a) et (b) opèrent comme des ressorts supplémentaires qui restreignent le mouvement des noeuds.

**Étape 3. Satisfaction des contraintes :** si la finalité est d'avoir une segmentation en régions, une procédure d'étiquetage en composantes connexes est exécutée sur le contour produit (qui n'est malheureusement pas toujours fermé). Une étape de fusion de régions peut être ensuite effectuée pour satisfaire certaines contraintes (des régions compactes et de grandes tailles par exemple). Cette étape est optionnelle et dépend du domaine d'application envisagé.

L'inconvénient de cette approche est que l'estimation initiale et le lissage dépendent d'un paramètre, qui est le rayon  $\rho$  du cercle sur lequel l'estimation est effectuée pour chaque point du contour. De plus, des contours fermés en entrée ne donnent pas forcément un contour fermé en sortie, ce qui pourrait compromettre le traitement ultérieur de ce contour par le processus de haut-niveau.

### 2.3.4.5 Coopération dynamique en boucle fermée

Une autre approche proposée par Bajcsy et al. [19] [44] exploite la complémentarité régions-contours et consiste à faire fonctionner deux algorithmes de segmentation alternativement en boucle fermée, jusqu'à ce que les résultats obtenus soient jugés satisfaisants.

Le système est composé de deux modules : un détecteur de contours et un détecteur de régions. Chaque module admet une image en entrée et fournit un résultat en sortie. Plusieurs paramètres de contrôle sont également fournis en entrée de chaque module afin de régler la formation des contours et des régions. En sortie, chaque module comporte des paramètres qui permettent de quantifier les résultats obtenus.

Le résultat final de la segmentation est obtenu après exécution d'un certain nombre de boucles sur chacun ou sur tous les modules en même temps. Le processus est itératif, les résultats de la segmentation obtenus lors d'une itération (quantifiés par les paramètres de sortie) sont utilisés pour adapter au mieux le ou les modules lors de l'itération suivante.

Chaque module peut fonctionner en boucle fermée sur lui même en ajustant ses propres paramètres d'entrée en fonction de ses sorties.

Dans le cas où le système global (c'est-à-dire les deux modules chaînés) fonctionne en boucle fermée, son résultat (image de régions) adapte les paramètres d'entrée de chaque module. L'algorithme de détection de contours utilisé est celui de Canny [60] qui présente l'avantage d'être paramétrable. La détection de région s'opère par un algorithme de croissance de régions.

Le module de détection de contours utilise deux paramètres en entrée: la taille de la fenêtre d'opérateur de lissage utilisée par le filtre de Canny, et le seuil sur la norme du gradient. En sortie, le détecteur de contours mesure deux paramètres: la densité de points de contours de l'image et la longueur moyenne des contours.

Pour le module d'extraction de régions, les paramètres d'entrée sont: le seuil sur la différence locale admise à l'intérieur d'une même région, la taille minimale de la région, et un paramètre de convexité. En sortie, deux attributs sont calculés: la taille moyenne des régions obtenues (ou le nombre) et le coefficient de compacité moyen de régions (donné par le rapport périmètre/surface).

Le paramètre de retour le plus important est le nombre de régions. En effet, le nombre de régions désirées est une donnée initiale. L'algorithme va modifier automatiquement ses paramètres de contrôle en entrée, à l'aide des rebouclages, de façon à ajuster au mieux le nombre de régions trouvées à l'issue du processus de segmentation. Cette procédure est ainsi itérée tant que le nombre de régions trouvées n'est pas proche de celui souhaité. En fait, tous les paramètres du système sont liés. Le nombre de régions dans l'image permet de déduire leur taille moyenne. L'estimation du rapport moyen entre le périmètre et l'aire des régions permet de donner une indication sur le facteur d'échelle du détecteur de contour, ainsi que la densité de points de contours dans l'image. Ce paramètre permet, moyennant une boucle de rétroaction sur le détecteur de contours, de fixer la valeur du seuil sur la norme du gradient. Ce dernier paramètre est lié au seuil sur la différence locale des niveaux de gris admise à l'intérieur d'une région. La taille minimale des régions est estimée par un pourcentage de la taille moyenne des régions désirées.

Cette méthode constitue une approche intéressante pour résoudre le problème de la segmentation. Toutefois, elle présente les inconvénients suivants: d'une part, l'introduction de connaissances a priori trop précises (le nombre de régions désirées et le rapport périmètre sur aire). En général, cette information n'est pas disponible, d'autre part, les règles précises de contrôle de la convergence du processus ne sont pas présentées.

## 2.4 Méthodes d'évaluation des résultats de segmentation

Le problème de la mesure de la qualité d'un résultat fourni par un algorithme de segmentation n'est toujours pas complètement résolu. L'établissement d'un critère de qualité objectif et absolu s'est avéré impossible devant la grande variété d'applications et l'incompatibilité des résultats avec le jugement subjectif de notre système visuel et vice-versa.

Il est très difficile de juger de la qualité d'une segmentation [93], car la segmentation n'est pas une fin en soi et ne possède un sens que dans le contexte d'une chaîne de traitement entière. Une bonne segmentation pourra alors être définie comme une segmentation donnant les éléments nécessaires et exploitables par les processus d'interprétation placés en aval. Mais cela ne veut-il pas dire que la segmentation d'une image dépend du domaine de l'application, c'est-à-dire, de ce qui est recherché dans cette image? Pour illustrer ce propos, considérons un exemple d'une image représentant un être humain. Si le but du traitement est de reconnaître cet être parmi une liste de personnes, une bonne segmentation va être celle qui fournira un maximum d'éléments et donc comprendra beaucoup de régions (oeil, bouche, détails du visage,...). Par contre, si le but est de localiser la position de la personne, la meilleure segmentation va être celle qui comporte uniquement deux régions: le fond et le corps. Cet exemple montre que bien qu'il s'agit d'une même image, plusieurs segmentations peuvent être considérées comme correctes, selon le contexte d'application choisi. Notons toutefois qu'en général, une sur-segmentation est préférable à une sous-segmentation. Celle-ci peut être facilement corrigée et affinée en procédant à une étape supplémentaire de regroupement, alors qu'une sous-segmentation est pénalisante pour la suite du traitement (interprétations,...).

Ceci étant, de nombreux chercheurs [10] [94] [105] [206] [252] [319] [342] [360] [368] [369] [370] se sont penchés sur ce problème et ont mis en oeuvre différentes approches d'évaluation de la segmentation.

Les différentes méthodes d'évaluation peuvent être classées en trois grandes catégories: les méthodes empiriques avec connaissance de la segmentation idéale; les méthodes empiriques sans connaissance de la segmentation idéale et les méthodes analytiques.

### 2.4.1 Evaluation avec connaissance de la vérité de terrain

Dans des applications pratiques de segmentation, des erreurs dans l'image segmentée sont tolérées. De fait, si les images à segmenter sont complexes et que les algorithmes sont entièrement automatiques, l'erreur est inévitable ou du moins fort probable.

L'erreur entre l'image segmentée et l'image de segmentation idéale (image de référence)

peut être le meilleur critère pour évaluer les performances des algorithmes. Selon le cas, l'image de référence peut être une image synthétique ou une image obtenue par des mesures sur le terrain (la vérité de terrain).

Les différents types de mesures basées sur la divergence proposés et utilisés sont les suivants:

- erreur basée sur le nombre de pixels mal segmentés;
- erreur basée sur la position de pixels mal segmentés;
- erreur basée sur le nombre d'objets dans l'image;
- erreur basée sur les valeurs des caractéristiques des objets segmentés;
- ...

Par exemple, Vinet [341], propose une mesure de la qualité de la segmentation qui consiste à rechercher les couples de régions les plus similaires dans les deux segmentations (la segmentation résultat et la segmentation de référence). La mesure de Vinet est alors proportionnelle au nombre de pixels n'intervenant pas dans cette superposition maximale. Cette mesure présente l'avantage de ne pas nécessiter de connaissance a priori sur la correspondance entre les régions d'une part, et, d'autre part, de ne pas exiger de segmentations comprenant le même nombre de régions. Elle présente, toutefois, l'inconvénient de ne pas tenir compte de l'information spatiale sur la localisation des pixels n'intervenant pas dans la superposition.

Kara-Falah et Bolon [182] définissent une nouvelle distance provenant de l'extension de la mesure de Baddeley [18]. Celle-ci tient compte de l'information spatiale concernant les pixels n'intervenant pas dans la superposition. Elle donne des résultats différents lorsque ces pixels proviennent de l'introduction de petites régions, ou lorsqu'ils proviennent de mauvaises détections de la frontière. Cette prise en compte d'informations spatiales permet, dans certains cas, de conduire à de meilleurs résultats que ceux obtenus par la mesure de Vinet.

Dans le cas d'une détection de contour, deux catégories de mesures peuvent être distinguées: celles relatives aux erreurs de localisation et celles sensibles aux erreurs de détection.

#### 2.4.1.1 Mesure de la performance de la localisation

Parmi les mesures les plus utilisées dans cette première catégorie se trouve l'indice de mérite défini par Pratt [283]. Soit  $\mathcal{R}$  l'image binaire de référence représentative de la vraie segmentation et soit  $\mathcal{S}$  l'image obtenue par la procédure de segmentation. Cette mesure est donnée par

$$Proba(\mathcal{R}, \mathcal{S}) = \frac{1}{\max\{n(\mathcal{R}), n(\mathcal{S})\}} \sum_{x \in \mathcal{S}} \frac{1}{1 + \alpha d(x, \mathcal{R})^2} \quad (2.124)$$

### 2.4.1.2 Mesure de la performance d'une détection

Cette mesure indique le taux de mauvaise classification des points de l'image. Les points appartenant à  $\mathcal{S}$  et pas à  $\mathcal{R}$  seront appelés points de type I. Les points qui appartiennent à  $\mathcal{R}$  mais pas à  $\mathcal{S}$  seront notés points de type II. Les mesures correspondant à ces deux types d'erreurs sont données par

$$\alpha(\mathcal{R}, \mathcal{S}) = \frac{n(\mathcal{S} \setminus \mathcal{R})}{n(I \setminus \mathcal{R})} \quad (2.125)$$

$$\beta(\mathcal{R}, \mathcal{S}) = \frac{n(\mathcal{R} \setminus \mathcal{S})}{n(\mathcal{R})} \quad (2.126)$$

où  $I$  désigne l'ensemble des points de l'image;  $n(\mathcal{R})$  est le nombre de point dans  $\mathcal{R}$ . L'opération  $(\mathcal{S} \setminus \mathcal{R})$  est équivalente à l'ensemble  $\{x \in \mathcal{S} / x \notin \mathcal{R}\}$

Une autre mesure employée est fondée sur le pourcentage de points mal classés. Ce type de mesure présente les inconvénients suivants:

- la non prise en compte de l'information spatiale inhérente aux points mal classés. En effet, un pixel situé au centre d'un objet est pondéré avec le même poids qu'un autre point situé sur le contour du même objet.
- l'absence d'information sur les classes des points responsables de l'erreur observée.

Yassnof et al. [367] proposent l'utilisation d'une matrice de confusion où les lignes correspondent aux classes réelles auxquelles les points appartiennent et les colonnes représentent les classes auxquelles les points ont été affectés. Les éléments  $C_{ii}$  de la diagonale de la matrice correspondent au nombre de points correctement étiquetés tandis que les termes  $C_{ij}$  représentent le nombre de points étiquetés  $i$  alors qu'ils appartiennent à la classe  $j$ . Pour chacune des classes deux mesures sont calculées. La première mesure indique une erreur de classification de type affectation d'une mauvaise étiquette (attribution à un point de la classe  $k$  une étiquette différente). La seconde mesure est représentative des points étiquetés avec une classe  $k$  alors qu'ils appartiennent à une autre classe.

## 2.4.2 Evaluation sans connaissance de la vérité de terrain

Les méthodes d'évaluation dans ce groupe utilisent des mesures dont la plupart sont définies de manière intuitive en se posant la question suivante: quelles conditions pourraient être satisfaites par une segmentation idéale?

Les différents algorithmes d'évaluation sont caractérisés simplement par le calcul des mesures définissant la segmentation idéale basée sur l'image segmentée sans connaissance de la vraie segmentation. Les différents types de mesures proposés pour caractériser la segmentation idéale sont les suivants *l'uniformité intra-région, le contraste inter-région et la forme des régions*.

Levine et Nazif, dans [215], définissent un ensemble de mesures, calculées à partir d'une segmentation d'images, qui rend compte de la différence entre la segmentation obtenue et celle qu'ils considèrent comme optimale. Ils définissent alors la segmentation optimale comme une partition de l'image en régions uniformes et ayant des frontières fortement contrastées. A partir de cette définition, un vecteur de performance est évalué pour différentes fenêtres d'analyse  $w$  de l'image. Ce vecteur comprend plusieurs mesures:

$$\vec{V}_w = [U_{w1}, \dots, U_{wm}, C_{w1}, \dots, C_{wm}, H_{w1}, \dots, H_{wm}, T_w, R_w, L_w] \quad (2.127)$$

où, pour une zone  $w$ :

$U_{wi}$ ,  $C_{wi}$  et  $H_{wi}$  ( $i = 1, \dots, m$ ) représentent des mesures d'uniformité, de contraste et d'énergie sur des lignes;

$T_w$  correspond à une mesure de connectivité entre les lignes;

$R_w$  et  $L_w$  sont des valeurs du nombre de régions et du nombre de lignes.

Ce vecteur de caractéristiques est utilisé par la suite pour guider de manière itérative la segmentation et donc, retravailler à l'intérieur de chaque région afin de rendre la segmentation la plus proche possible de la segmentation idéale. Ceci est réalisé grâce à un superviseur qui décide des différentes règles à appliquer aux régions (division, fusion,...).

Dans [368], Zhang et Gerbrands considèrent que le but final de la segmentation (et des autres traitements) est souvent d'obtenir des mesures des caractéristiques des objets de la scène. En conséquence, la précision de ces mesures, calculées à partir de l'image segmentée, peut être un bon élément pour juger de la performance d'un algorithme de segmentation. Ils proposent donc de mesurer les caractéristiques des objets segmentés telles que le périmètre, la taille, la circularité,....

A partir de ces mesures et connaissant la nature de l'objet (image de synthèse), une mesure de la qualité de la segmentation peut être faite. De nombreux problèmes apparaissent en utilisant cette méthode. Tout d'abord, elle implique l'utilisation d'images de synthèse bien connues, et donc rejette toutes les images issues de scènes courantes. Un autre problème apparaît: cette méthode ne peut pas s'appliquer si l'objet a été segmenté en plusieurs régions, ce qui est souvent le cas avec des images bruitées.

La principale critique concernant ces travaux est liée au fait que toutes les mesures effectuées sont issues d'une définition générale de la segmentation idéale. Or la segmentation dépend du contexte dans lequel elle est réalisée (de ce qui est recherché dans l'image), ce qui n'est pas du tout pris en compte ici. La segmentation considérée comme optimale par rapport à toutes les mesures déterminées, n'est donc pas forcément celle évaluée par un observateur humain dans un contexte précis.

### 2.4.3 Méthodes analytiques d'évaluation

Ces méthodes examinent et évaluent les algorithmes de segmentation, en analysant leurs principes et leurs propriétés. Ces méthodes évitent l'exécution concrète de l'algorithme et le résultat est donc indépendant des images à segmenter.

Jusqu'à maintenant, ces méthodes travaillent sur des modèles particuliers ou sur des propriétés des algorithmes. Les critères pris en compte sont par exemple, le type et le nombre de connaissances a priori incorporées dans l'algorithme, un seuil de probabilité dans une procédure statistique ou la stratégie et la complexité des calculs. Cependant, toutes les propriétés de l'algorithme de segmentation ne peuvent pas être obtenues par l'étude analytique. La difficulté réside dans le manque de cadre théorique général sur la segmentation d'images.

Dans leur revue des différentes techniques d'évaluation de méthodes de segmentation [370], Zhang et al. posent le problème de la quantification de la qualité d'une segmentation par une approche analytique.

## 2.5 Conclusion

L'ensemble des techniques présentées dans les sections antérieures montrent que l'approche coopérative constitue une voie de recherche très intéressante et mérite une étude approfondie en vue de résoudre différents problèmes mal posés dans le domaine de la vision artificielle.

Avant de présenter, dans le chapitre suivant, notre approche de segmentation d'images, nous rappelons les principales remarques tirées des développements précédents.

- Nous avons vu dans la section 2.2.1 que la segmentation dans l'espace paramétrique consiste à regrouper les pixels de niveaux de gris semblables, indépendamment des relations de connexité qui les lient. Par exemple, le seuillage d'histogramme s'appuie sur l'hypothèse que les régions de niveau uniforme produisent des pics suffisamment significatifs pour que l'on puisse les caractériser directement par les valeurs limites des pixels qui les composent. Il suffit alors de seuiller l'image entre ces deux limites pour en extraire la région. L'inconvénient majeur de ces méthodes est qu'elles effectuent la segmentation sans prendre en compte la topologie de l'image.
- La propriété de la topologie est considérée dans les techniques par croissance de régions. En plus des caractéristiques photométriques des points de l'image, leur dimension spatiale (coordonnées) est prise en considération dans le processus de formation des régions. Nous avons toutefois constaté que l'application de ces méthodes peut être

autres introduire trois types d'erreurs :

- une frontière entre régions n'est pas un vrai contour et il n'y a pas de contours dans sa proximité;
- une frontière correspond à un contour mais ne coïncide pas avec celui-ci;
- l'existence des contours sans qu'il y ait des frontières correspondantes.

La principale raison qui engendre les deux premiers type d'erreur (production de fausses frontières ainsi que leur mauvaise localisation) est due à la définition stricte de l'uniformité de la région. Par exemple, si l'intensité lumineuse varie linéairement à l'intérieur d'une région  $R$  et que l'on impose à l'intensité d'être plus ou moins constante dans toute la région, alors il y aura production de fausses frontières. Le dernier type d'erreur est principalement dû aux bruits superposés sur les contours et ayant des niveaux de gris similaires à la région en cours de formation ce qui entraîne une sous-segmentation.

Notons enfin que les méthodes de croissance de régions sont en général adaptées aux zones uniformes de l'image. Les autres types de régions telles que les textures ou les dégradés ou encore les textures dégradées sont rarement considérés, sinon ils sont largement sur-segmentés.

- Nous avons également vu dans la section 2.2.2 que la détection de contours est une méthode généralement basée sur la détection locale des variations de niveau du signal image. Elle consiste classiquement à appliquer un opérateur gradient (resp. laplacien) puis un algorithme de recherche d'extrema (resp. des passages par zéro). Les méthodes de détection de contours supposent généralement une connaissance a priori d'un modèle à détecter. Elles fournissent souvent des frontières de régions discontinues. Un post-traitement est nécessaire lorsque l'on souhaite aussi connaître les pixels qui composent la région. Par ailleurs, elles sont sensibles aux variations spatiales brusques de l'image, qu'il s'agisse de bruit ou de texture, ce qui contribue à la production de faux contours à l'intérieur des régions.
- Dans la section 2.2.3, nous avons étudié les méthodes d'extraction de textures. Du fait de la nature contextuelle et globale de l'analyse, les méthodes de segmentation de textures souffrent dans leur majorité du problème de mauvaise localisation des frontières des régions. Elles présentent en outre le défaut d'altérer les structures fines dans l'image qui sont souvent soit fusionnées aux régions voisines soit sur-segmentées.
- Dans la section 2.3, nous avons exposé quelques méthodes coopératives en les classant en trois familles (séquentielles, parallèles, hybrides).

Les différentes approches de la coopération séquentielle conduisent à la conception d'algorithmes de segmentation puissants. Cependant, elles souffrent de faibles performances dues à l'ordonnancement imposé dans la coopération. De plus, dans la plupart des méthodes, le niveau d'intervention des régions est différent de celui des contours.

En ce qui concerne les approches parallèles, et selon les cas, l'accent est plutôt mis sur les problèmes de fusion, d'adaptation, ou de correction. La notion de coopération active n'intervient pas explicitement. Cette notion implique l'introduction simultanée des trois mécanismes sus-cités (fusion, adaptation, correction). Par exemple, dans le cas de la coopération contours-régions, elle sous-entend d'une part, le calcul des régions ou de contours, en fonction des attributs localement disponibles et, d'autre part, la remise en question bilatérale entre régions et contours, par introduction de mécanismes de corrections à l'aide de boucles de rétroaction.

Les approches de la coopération hybrides permettent de pallier à quelques lacunes des méthodes des deux précédentes catégories, elles présentent néanmoins le défaut lié au manque de critères décisionnels permettant la validation de la cohérence des indices extraits.

# Chapitre 3

## Systeme de segmentation proposé

L'objectif de ce chapitre est de présenter l'approche coopérative de segmentation d'images que nous avons développée. Cette approche repose sur un constat de relative inefficacité des techniques classiques de segmentation prises chacune séparément dans le cas des images riches en détails et présentant des régions de nature différente. Elle est fondée sur la combinaison de plusieurs techniques de segmentation. Une importance particulière a été accordée à l'aspect automatique et non-supervisé de l'approche.

Des résultats sur plusieurs types d'images seront présentés et des études comparatives permettront alors de tirer quelques conclusions.

### 3.1 Introduction et principe général

Pour analyser des images réelles riches en détails et contenant une grande variété de régions de nature uniforme et/ou texturée, il s'avère nécessaire d'adopter une démarche où la formation des primitives s'opère de manière coopérative et guidée. Une coopération fiable doit réaliser une efficace exploitation des résultats issus de chacune des méthodes impliquées: chaque méthode doit pouvoir améliorer la qualité de ses résultats, à la suite de l'information dont elle peut bénéficier des autres méthodes. L'efficacité de la coopération peut aussi être améliorée en introduisant, en plus des mécanismes de focalisation, de correction et d'adaptation aux contextes examinés, des critères décisionnels permettant d'évaluer le degré de robustesse des primitives extraites par les différentes techniques impliquées dans la coopération.

C'est dans cette perspective que se situe notre travail et nous proposons une architecture d'un système de segmentation faisant coopérer plusieurs techniques de segmentation, et intégrant un mécanisme de rétroaction pour permettre la remise en cause des résultats obtenus et l'ajustement des paramètres de contrôle des différentes techniques de segmentation. L'idée repose sur le fait qu'un résultat commun à plusieurs techniques de segmentation peut être considéré comme significatif. Connaissant les faiblesses de chacune des différentes

approches, il s'agit aussi d'exploiter au mieux les différents types d'informations pour aider à la formation de chacune des primitives dans les cas problématiques. En outre, afin d'engendrer une représentation cohérente et stable de l'image, un mécanisme pour vérifier la cohérence des primitives extraites par comparaison entre les différentes segmentations, est introduit.

Le principe général consiste à lancer les différentes techniques du système avec des critères stricts pour l'initialiser avec des résultats de sur-segmentation. Ensuite, les différents processus de segmentation sont itérés en relâchant certaines contraintes jusqu'à la convergence vers des résultats compatibles et stables. La formation de chaque type de primitives est effectuée en utilisant les résultats issus des autres segmentations. A la fin de chaque itération, de nouvelles primitives sont déterminées et des affinements des résultats de segmentations sont effectués. Une transformation des différents résultats permet de définir un mode de représentation commun des informations à mettre en correspondance. La représentation par contours a été adoptée. Une mesure de cohérence entre les différents résultats transformés permet de décider de la continuité ou de l'arrêt de la procédure globale de segmentation. La mesure de cohérence est effectuée à l'aide d'une mesure prenant à la fois en compte la détection des contours et leur localisation. Dans le cas où le critère d'arrêt n'est pas vérifié, les paramètres qui contrôlent la formation des différentes primitives sont réactualisés et la procédure est itérée. Le résultat de la segmentation qui sera retenu est celui qui maximise la cohérence entre les résultats.

Le système développé est composé de deux modules. Le *premier* est dédié à l'extraction de régions uniformes ou faiblement texturées. Il est constitué de deux méthodes : détecteur de contours et extracteur de régions par agrégation de points. Ces deux méthodes coopèrent mutuellement de façon itérative, l'une utilisant les informations fournies par l'autre. Le *second* module fait intervenir les primitives "textures" afin de réactualiser et corriger les primitives contours et régions non significatives extraites au fur et à mesure des différentes itérations du premier module.

Dans la suite, nous commençons par décrire le premier module du système développé ainsi que les méthodes utilisées (cas de figure de la coopération contours-régions uniformes). Ensuite, nous présentons le module chargé de la segmentation au sens de la texture. Enfin, nous étudions dans le cas de la coopération contours-régions uniformes-textures, et analysons les résultats de segmentation obtenus en les comparant à d'autres approches.

## 3.2 Extraction de régions uniformes et/ou faiblement texturées

Le premier module fait intervenir la coopération de deux méthodes : la détection de contours et l'extraction de régions uniformes. Celles-ci présentent des caractéristiques qui

sont complémentaires sur bien des points, en particulier la précision de la localisation des contours, la résistance au bruit et la fermeture naturelle des frontières des régions. La fusion des informations duales s'effectue par l'extraction d'un type d'information guidée et/ou corrigée par l'information duale. Le principe du système dans ce cas de figure consiste à faire évoluer la formation des contours et régions, à partir de sur-détection de contours et de sur-segmentation en régions, en relâchant les critères de formation des régions qui est contrôlée par les contours. Ceux-ci seront par la suite réactualisés et corrigés en s'aidant des régions extraites au fur et à mesure des différentes itérations. Le processus de segmentation est itéré jusqu'à stabilité des résultats. A la fin de chaque itération, la carte des régions est transformée en une carte de contours. Une mesure de similarité entre celle-ci et le résultat de la détection de contours est effectuée pour mesurer la cohérence des deux résultats. Cette mesure tient à la fois compte de la détection et de la localisation des contours. Le résultat de segmentation retenu est celui qui maximise la cohérence entre les résultats, c'est-à-dire qui minimise une dissimilarité entre ces deux résultats.

Nous présentons d'abord les deux méthodes qui seront utilisées par le premier module. Ensuite, nous étudions l'interaction entre ces deux méthodes. Enfin, des résultats sur plusieurs types d'images sont présentés et permettront alors de tirer quelques conclusions.

### 3.2.1 Approche adoptée pour la détection des contours

A l'issue de l'étude préliminaire faite dans la section 2.2.2 et vu les remarques sur les différents détecteurs de contours, nous avons choisi d'utiliser l'approche de Deriche. En effet, en plus de son optimalité au sens des critères de Canny (cf. paragraphe 2.2.2.4.1), le filtre développé par cette approche présente la propriété d'être paramétrable par un facteur d'échelle  $\alpha$  qui contrôle le degré de lissage.

L'étude du comportement du filtre à différentes échelles montre que son application pour des valeurs élevées de  $\alpha$  (basses échelles) permet une meilleure localisation de contours au détriment d'une forte sensibilité aux bruits. Le choix d'une échelle optimale reste un problème ouvert et une sur-segmentation étant préférable à une sous-segmentation, une détection de contours avec une valeur élevée s'est avérée préférable. En moyenne  $\alpha$  varie entre 0.5 et 2.

Pour des raisons de temps de calcul, l'implémentation récursive de l'opérateur de Canny (filtre de Deriche) a été adoptée (cf. annexe A).

Pour illustrer le comportement de l'opérateur de Deriche, nous présentons quelques résultats obtenus sur un contour 1D de type marche bruité et sur l'image SPOT "Aquitaine" issue de la banque d'images du GdR PRC-ISIS.

Les figures 3.2 et 3.3 illustrent le comportement de l'opérateur première dérivée en présence d'un contour 1D de type marche d'amplitude bruité présentant un rapport signal

à bruit SNR respectivement de 1 et 0.25 (Fig. 3.1) [108].

Les résultats correspondant à diverses valeurs de  $\alpha$  ont été obtenus avec seulement 5 opérations par point par la mise en œuvre récursive à une dimension décrite dans l'annexe A.

FIG. 3.1 - Contour de type "marche" bruité. SNR=1 à gauche et SNR=0.25 à droite [108].

FIG. 3.2 - Filtrage de la "marche" bruitée avec un SNR=1.  $\alpha=.25$ ;  $\alpha=.5$ ;  $\alpha=.75$ ;  $\alpha=1$ .

FIG. 3.3 - Filtrage de la "marche" bruitée avec un SNR=0.25.  $\alpha=.25$ ;  $\alpha=.5$ ;  $\alpha=.75$ ;  $\alpha=1$ .

Ces figures illustrent parfaitement le rôle du paramètre  $\alpha$ . Il doit être choisi assez grand (de l'ordre de 1 et plus) si on veut favoriser la localisation au détriment de la détection (pour

les signaux très peu bruités) et assez petit (de l'ordre de 0.25 à 0.5) si on veut privilégier la détection dans le cas de signaux très bruités.

Les images de la figure 3.4 montrent des exemples de contours extraits à partir de l'image "Aquitaine" présentée par la figure 3.4(a), en utilisant différentes valeurs de  $\alpha$  et en mettant tous les maxima locaux non-nuls à zéro (noir) et le reste à 255 (blanc). Nous pouvons, également dans ce cas, constater que la diminution de la valeur de  $\alpha$  favorise la bonne détection au détriment de la localisation et vice versa. En effet, pour des hautes fréquences (resp. basses fréquences)  $e^{-|\alpha|}$  tend vers 0 (resp. vers 1) et par conséquent  $\alpha$  est grand (resp.  $\alpha$  est proche de 0).

(a)

(b)

(c)

(d)

FIG. 3.4 - (a) Image SPOT de la région Aquitaine après réhaussement. Détection des maxima locaux par filtre de Deriche. (b)  $\alpha=0.5$ ; (c)  $\alpha=1$ . (d)  $\alpha=2$ . Les points en noir correspondent aux maxima des gradients non nuls.

### 3.2.2 Approche adoptée pour l'extraction de régions uniformes

L'objet de cette section est de présenter les méthodes de segmentation d'images en régions que nous avons développées. Le but est de fournir une représentation pertinente de l'image avec des primitives de type "régions uniformes" au sens des niveaux de gris.

Nous avons vu (section 2.2.1.3) que la formation de régions peut être réalisée de deux manières : par fusion et/ou division d'ensemble de points fondée sur des critères globaux ou par agrégation de points basée sur des critères de similarité locaux. Il nous semble que l'appartenance d'un pixel à une région nécessite la prise en compte d'informations sur la région et d'informations locales relatives au pixel. Or, dans un processus de fusion, puisque l'étude se fait au niveau des régions, on perd l'information locale.

L'approche de croissance de régions par agrégation de points, en utilisant des critères qui tiennent compte à la fois du contexte local et des informations sur la région formée, a donc été retenue. Cette approche permet une meilleure efficacité quant à la rapidité et à la taille de mémoire nécessaire à la structure des données relatives aux régions et à leurs attributs.

Dans les paragraphes suivants, nous présentons les algorithmes de croissance de régions, les structures de données ainsi que les modèles utilisés pour représenter l'image et les régions.

#### 3.2.2.1 Présentation des algorithmes de croissance de régions

Les différents algorithmes de croissance de régions reposent sur des principes de fonctionnement identiques :

- définition d'un prédicat  $P_{rp}(R_c, p_c)$ , pour mesurer le degré de similarité du pixel examiné  $p_c$  avec la région en cours de formation  $R_c$ ;
- initialisation de la région et de son modèle par un ensemble de pixels, noté  $G$  et appelé noyau ou région-noyau;
- croissance de la région en respectant la connexité des pixels avec la région et en examinant la possibilité d'appartenance à la région en cours de formation.

Nous noterons  $p_c$  le pixel courant,  $p_v$  un pixel voisin du pixel courant au sens de la 8-connexité, et  $R_c$  la région courante.

Le corps de l'algorithme de croissance de régions est le suivant :

**Début algorithme** *croissance de régions*

**Tant qu'il existe des pixels non traités :**

$R_c := 0 + G$  ; Initialiser la région avec le noyau  $G$  ;

$F \rightarrow G$  ; Initialiser la file des points avec le noyau ;

**Appel** *fct-de-croissance* ( $R_c, p_c$ ) ;

**Fin tant que**

Afin de mettre en évidence les avantages et inconvénients des méthodes d'agrégation de points, plusieurs variantes de *fct – de – croissance* ont été examinées. Le premier algorithme de croissance de régions que nous présentons serait l'algorithme idéal, le second est un algorithme récursif, le troisième est une version améliorée du précédent. Quant au dernier, il représente un version isotrope assurant une croissance de régions spatialement homogène.

### 3.2.2.1.1 Algorithme idéal

La fonction de croissance de régions idéale serait la suivante :

```

Fonction fct-de-croissance-idéale (région  $R_c$  , pixel  $p_c$ )
  Pour tout  $p_v$ , voisin de  $p_c$  non traité
    Début séquence parallèle
      Si  $P_{rp}(R_c, p_v)$ 
         $R_c := R_c + p_v$ ; Ajouter  $p_v$  à la région; mettre à jour  $R_c$ ;
         $p_c := p_v$ ;
        fct-de-croissance-réursive-idéale ( $R_c, p_c$ );
      Fin si
    Fin séquence parallèle
  Fin pour tout
Fin Fonction

```

Dans cet algorithme, nous remarquons l'apparition d'une séquence qui s'exécute en parallèle et qui pose des problèmes de mise en œuvre logicielle sur des machines mono-processeurs. Nous proposons dans l'algorithme suivant une approximation permettant de supprimer la séquence parallèle de la fonction de croissance.

### 3.2.2.1.2 Algorithme récursif

En supprimant la séquence parallèle de l'algorithme idéal, celui-ci devient :

```

Fonction fct-de-croissance-réursive (région  $R_c$  , pixel  $p_c$ )
  Pour tout  $p_v$ , voisin de  $p_c$  non traité
    Si  $P_{rp}(R_c, p_v)$ 
       $R_c := R_c + p_v$ ; Ajouter  $p_v$  à la région; mettre à jour  $R_c$ ;
       $p_c := p_v$ ;
      fct-de-croissance-réursive ( $R_c, p_c$ );
    Fin si
  Fin pour tout
Fin Fonction

```

Le principal avantage de l'algorithme récursif réside dans sa simplicité d'implémentation mais il présente l'inconvénient de ne tenir compte d'aucune hiérarchie dans l'appartenance des points examinés.

Les points ayant un fort degré d'appartenance à la région en cours de formation sont traités de la même façon que les autres points. Ce problème est évité en effectuant un test sur le degré de similarité du point à la région : la priorité d'annexion est donnée au point qui maximise le degré de similarité. Ainsi, la formation de la région est optimale au sens de la similarité entre points qui la composent.

### 3.2.2.1.3 Algorithme récursif optimal

La fonction de croissance récursive est modifiée de la manière suivante :

**Fonction** *fct-de-croissance-réursive-optimale* (région  $R_c$  , pixel  $p_c$ )  
**Pour tout**  $p_v$ , voisin de  $p_c$  non traité  
  **Si**  $P_{rp}(R_c, p_v)$   
    **Choisir**  $p_v$  tel que  $P_{rp}(R_c, p_v)$  soit optimal;  
     $R_c := R_c + p_v$ ; Ajouter  $p_v$  à la région; mettre à jour  $R_c$ ;  
     $p_c := p_v$ ;  
    *fct-de-croissance-réursive-optimale* ( $R_c, p_c$ );  
  **Fin si**  
**Fin pour tout**  
**Fin Fonction**

Dans ce cas, la formation de la région n'est pas conditionnée par une direction donnée, mais dépend uniquement des données. Toutefois, ce type de croissance s'avère inefficace lorsque le noyau qui caractérise la région n'est pas fiable car dans ce cas l'algorithme converge vers la meilleure région au sens de la ressemblance au noyau.

La croissance des régions est non homogène dans le sens où des points éloignés du noyau seront testés et éventuellement annexés alors que des points plus proches n'auront pas été testés.

### 3.2.2.1.4 Algorithme isotrope

Afin de remédier au dernier problème, un contrôle sur le choix des voisins du point examiné est introduit de manière à tenir compte de l'éloignement du pixel testé par rapport au noyau.

Cette idée est implémentée en mémorisant les frontières de la région en cours de formation dans une file  $F$ . Les points de la frontière sont traités séquentiellement. Lorsqu'un nouveau point est annexé, la frontière est mise à jour, les nouveaux points de celle-ci sont placés dans la file et sont donc traités après les points de l'ancienne frontière.

L'algorithme de croissance de régions sera dans ce cas de figure le suivant :

```

Tant qu'il existe des pixels non traités : choisir un noyau  $G$ 
   $R_c := 0 + G$  ; Initialiser la région à partir de  $G$  ;
   $G \rightarrow F$  ; Initialiser la file avec le noyau ;
  Tant que  $F$  n'est pas vide
     $F \rightarrow c$  ; Prendre le premier élément de la file ;
    Pour tout  $p_v$ , voisin de  $p_c$  non traité
      Si  $P_{rp}(R_c, p_v)$  est optimal ;
         $R_c := R_c + p_v$  ; Ajouter  $p_v$  à la région ; mettre à jour  $R_c$  ;
         $v \rightarrow F$  ; Ajouter  $p_v$  dans la file ;
      Fin si
    Fin pour tout
  Fin tant que
Fin tant que

```

### 3.2.2.2 Choix d'un algorithme de croissance de régions

L'algorithme récursif présente l'avantage d'être simple à implémenter mais aucune hiérarchie dans l'appartenance des points examinés n'est prise en compte. Ce problème est évité en donnant une priorité d'annexion au point qui maximise le degré de similarité.

Dans le cas de l'algorithme récursif optimal, la formation de la région ne dépend pas d'une direction imposée a priori, mais des données. Cette croissance s'avère néanmoins non satisfaisant quand le noyau n'est pas fiable car dans ce cas l'algorithme converge vers la meilleure région au sens de la ressemblance au noyau. Des points éloignés du noyau seront testés et éventuellement agrégés alors que des points plus proches n'auront pas été examinés. Afin de pallier ce problème, un test sur l'éloignement du pixel par rapport au noyau est ajouté. Une croissance isotrope des régions est garantie (indépendante d'une direction donnée) mais nécessite plus de temps que l'algorithme récursif.

Notons que pour un jeu de paramètres restrictif (faibles seuils de segmentation), il n'y a pas de différence notable entre les segmentations fournies par les différents algorithmes. Étant donnée notre stratégie d'initialisation avec des résultats de sur-segmentation et pour des raisons de temps de calcul, nous avons opté pour l'algorithme récursif. Ce processus de croissance de régions est simple à mettre en œuvre. Toutefois, en plus d'une localisation imprécise des frontières, il engendre, comme c'est le cas des autres variantes d'agrégation de points, des sous-segmentations malgré l'utilisation de faibles seuils d'agrégation. La principale raison qui engendre la mauvaise localisation des frontières est due à la définition stricte de l'uniformité de la région. Les erreurs relatives aux sous-segmentations sont principalement dues aux bruits superposés sur les frontières et ayant des niveaux de gris similaires à la région en cours de formation.

Dans les différents algorithmes présentés, une fonction de décision (critère de similarité)  $P_{rp}(R_c, p_c)$  ainsi que des régions-noyaux et des modèles des régions doivent être définis. L'objet des paragraphes suivants est de présenter l'ensemble de ces différents éléments.

### 3.2.2.3 Critère de similarité

Cette fonction mesure le degré de similarité de l'attribut (l'intensité  $I(i, j)$ ) du pixel examiné  $p_c$  de coordonnées  $(i, j)$  avec le modèle de la région  $R_c$  en cours de formation. Comme il a été mentionné au paragraphe 2.2.1.3.2, une combinaison de plusieurs critères peut être utilisée. Dans notre cas, le test de Fisher est employé. Celui-ci donne une indication sur la similarité entre l'intensité  $I(i, j)$  du pixel examiné et la moyenne  $\mu_{R_c}$  de la région en cours de formation. Il est défini par la relation suivante :

$$P_{rp}(R_c, p_c) \equiv \left[ \frac{|\mu_{R_c} - I(i, j)|}{\sigma_{R_c}} \leq S_k \right] \quad (3.1)$$

où  $\sigma_{R_c}$  représente l'écart-type de la région  $R_c$  et  $S_k$  est une valeur de seuil.

Notons que dans le processus d'agrégation, il peut exister des ensembles de pixels qui ne vérifient aucun critère de groupement. De tels pixels, souvent situés sur une discontinuité, ou même en bordure de celle-ci, ont souvent des caractéristiques locales qui ne sont plus celles de la région voisine. L'annexion de tels pixels entraînerait alors une variation arbitraire de la moyenne et de l'écart-type des niveaux de gris de la région. Il est alors souhaitable de les classer dans une catégorie de pixels problématique et de retarder la décision de leur annexion à telle ou telle région jusqu'à ce que l'on puisse obtenir des informations issues d'autres traitements (contours, textures,...).

### 3.2.2.4 Détermination des régions-noyaux

Généralement les noyaux des régions sont introduits manuellement par un opérateur. Dans le cas non supervisé, les régions fiables de l'image peuvent être déterminées en effectuant des opérations logiques entre plusieurs cartes de régions différentes pour la même image [184]. Les noyaux des régions sont alors obtenus par intersection de plusieurs résultats de segmentations, réalisées par la même procédure de croissance de régions, mais en utilisant différentes directions de parcours des pixels et en appliquant des seuils de segmentation faibles pour garantir la stationnarité de l'homogénéité des régions (figure 3.5).

FIG. 3.5 - *Extraction d'un noyau homogène.*

Une autre approche basée sur l'analyse de l'histogramme global qui sera présentée dans la section 3.3.1.1 peut également être utilisée. En supposant la multi-modalité de l'histogramme et une bonne séparation de ses différents modes, le principe est de détecter des noyaux localisés à l'intérieur des régions homogènes de l'image. Ceux-ci correspondent aux pixels ayant les niveaux de gris des pics de l'histogramme. Pour avoir des noyaux homogènes, un seuillage plus restrictif (les seuils ne sont pas placés aux vallées mais à des niveaux intermédiaires entre les pics et les vallées) est nécessaire.

### 3.2.2.5 Modèles et structures des données

Deux types d'attributs sont nécessaires dans le modèle de région : les attributs descriptifs et les attributs internes. Les attributs internes sont des grandeurs utilisées par le processus de segmentation. Souvent, les attributs (statistiques et/ou géométriques) sont à la fois descriptifs et internes. Le vecteur de caractéristiques de la région est constitué du nombre de pixels de la région, noté  $N_{R_k}$ , de l'intensité moyenne et de l'écart-type des niveaux de gris de ses pixels.

La représentation utilisée pour modéliser les régions est composée d'un graphe d'adjacence afin de les traiter comme des entités. Chaque noeud du graphe représente une région et son vecteur caractéristique, tandis que les arcs représentent l'adjacence entre les différentes régions.

#### 3.2.2.5.1 Initialisation et mise à jour de la région

Lors du démarrage du processus de croissance de régions, le modèle de la région courante est initialisé en fonction des paramètres du noyau.

Au cours de la croissance de régions, lorsqu'un pixel est ajouté à une région, le modèle de la région est mis à jour. La mise à jour des paramètres d'un modèle, lorsqu'une nouvelle observation est disponible, est un problème classique d'estimation. Deux approches sont possibles selon que l'estimateur prend une forme récursive ou non. Un estimateur récursif peut s'écrire :

$$v^{k+1} = F(v^k, i^{k+1}) \quad (3.2)$$

où  $v^k$  est un vecteur de paramètres estimé pour  $k$  informations et  $i^{k+1}$  est l'innovation apportée par la  $(k + 1)$ ème donnée.

Avec ce type d'estimateur, et contrairement à l'estimateur non récursif, il n'est pas nécessaire de manipuler toutes les données pour calculer la nouvelle valeur des paramètres. De ce fait, les estimateurs récursifs sont intéressants en segmentation d'images où les régions peuvent contenir un nombre considérable de points. Ils permettront une réduction du temps de traitement. C'est pour cette raison que nous avons choisi d'utiliser des estimateurs récursifs pour calculer les paramètres du modèle de région.

### 3.2.2.5.2 Estimation récursive des paramètres de la région

Les estimateurs empiriques de la moyenne et de l'énergie peuvent prendre une forme récursive selon la façon suivante.

Soient  $x$  une variable à estimer et  $x_k$  les observations. La moyenne et l'énergie pour les  $n$  premières observations peuvent respectivement être estimées par :

$$\hat{\mu}_{x_{(n+1)}} = \frac{1}{n+1} \sum_{k=1}^{n+1} x_k = \frac{n}{n+1} \hat{\mu}_{x_{(n)}} + \frac{1}{n+1} x_{(n+1)} \quad (3.3)$$

$$\hat{\sigma}_{x_{(n+1)}}^2 = \frac{1}{n+1} \sum_{k=1}^{n+1} x_k^2 = \frac{n}{n+1} \hat{\sigma}_{x_{(n)}}^2 + \frac{1}{n+1} x_{(n+1)}^2 \quad (3.4)$$

Notons que l'estimateur de variance s'écrit simplement en fonction des deux paramètres estimés précédemment :

$$\hat{\sigma}_{(n+1)}^2 = \hat{\sigma}_{x_{(n+1)}}^2 - (\hat{\mu}_{x_{(n+1)}})^2 \quad (3.5)$$

En combinant les équations précédentes on obtient la forme récursive :

$$\hat{\sigma}_{(n+1)}^2 = \frac{n}{n+1} \left( \hat{\sigma}_{(n)}^2 + (\hat{\mu}_{x_{(n)}})^2 \right) + \frac{1}{n+1} x_{(n+1)}^2 - \left( \frac{n}{n+1} \hat{\mu}_{x_{(n)}} + \frac{1}{n+1} x_{(n+1)} \right)^2 \quad (3.6)$$

que nous réécrivons sous la forme de l'estimateur récursif d'écart type :

$$\hat{\sigma}_{(n+1)} \equiv F_{\sigma}(\hat{\sigma}_{(n)}, x_{(n+1)}) \quad (3.7)$$

Dans le paragraphe suivant nous présentons quelques résultats obtenus par l'application de l'algorithme récursif de croissance de régions sur l'image Aquitaine.

### 3.2.2.6 Résultats expérimentaux et conclusions

Pour homogénéiser les régions de l'image et régulariser les données en évitant d'altérer leur caractère texturé, un lissage adaptatif de l'image basée sur le principe de la diffusion anisotropique (cf. Annexe B) est effectué avant d'appliquer la procédure de croissance de régions.

Les images de la figure 3.6 montrent des exemples de régions obtenues à partir de l'image de la figure 3.4(a), en appliquant l'algorithme récursif après sélection des régions-noyaux par intersection de résultats de segmentations réalisées dans quatre directions (balayage de l'image ligne par ligne du haut en bas et du bas vers le haut, et balayage colonne par colonne de gauche à droite et de droite à gauche) et en utilisant différentes valeurs  $S_k$  du critère de similarité. Les résultats montrent que malgré l'utilisation de faibles seuils de similarité certaines régions de l'image sont sous-segmentées. On peut également remarquer la mauvaise localisation de la plupart des frontières des régions extraites.

Les méthodes qui tentent d'extraire les régions significatives de l'image supposent que les régions intéressantes sont homogènes. Le but des différentes techniques est souvent de partitionner l'image en respectant au mieux une ou plusieurs propriétés arbitraires, mais lors de l'application, ces propriétés sont fortement liées à l'homogénéité des régions. Cette hypothèse, un peu restrictive, de définition stricte de l'uniformité de régions est à la base d'un des principaux défauts des algorithmes de segmentation par extraction de régions, qui est la difficulté à positionner correctement les frontières des régions. Ainsi une région à l'intérieur de laquelle la luminosité varie linéairement peut être divisée en deux sous régions et deux régions distinctes ayant des contrastes faibles le long de leurs frontières peuvent être fusionnées en une seule. C'est une des raisons qui rendent nécessaire le recours à une information tirée de la détection de contours pour assister la technique de détection de régions.

(a)

(b)

(c)

(d)

FIG. 3.6 - *Extraction de régions par agrégation récursive de points appliquée à l'image Aquitaine. (a)  $S_k=6$ . (b)  $S_k=7$ . (c)  $S_k=8$ . (d)  $S_k=9$ .*

### 3.2.3 Coopération contours-régions uniformes

Après avoir présenté les méthodes de détection de contours et d'extraction de régions uniformes, nous allons, dans cette section, étudier la coopération et l'interaction contours-régions uniformes, sans intégration de l'information texture.

Idéalement, la description d'une image à partir des primitives contours et régions devrait être identique :

- un contour fermé définit une région,
- une région définit un contour fermé.

En pratique, les différences sont importantes et on obtient rarement des descriptions équivalentes à partir de ces primitives.

Comme il a déjà été signalé, typiquement, le principal avantage des contours est qu'ils sont localisés de manière précise. Néanmoins, l'application d'une approche contours se heurte souvent au problème de sous-détection de certaines discontinuités ce qui engendre des contours ouverts. Le point fort de l'approche par extraction de régions est justement la fermeture de ces frontières et la richesse de l'information qu'elles véhiculent. Leur localisation exacte reste cependant difficile à obtenir.

Le système de segmentation développé permet d'exploiter la dualité contour-région évoquée précédemment. Le synoptique général du système dans cette première approche peut être illustré par la figure 3.7. Il consiste à faire évoluer, à partir de sur-détection de contours et de sur-segmentation en régions, les deux procédures en relâchant les critères de formation des régions qui sont contrôlées par les contours (boucle (1) sur Fig. 3.7). Ceux-ci seront par la suite réactualisés et corrigés en s'aidant des régions extraites au fur et à mesure des différentes itérations (boucle (2) sur Fig. 3.7).

La détection des contours s'opère à l'aide d'un filtre exponentiel optimal au sens des critères de Canny-Deriche (paragraphe 2.2.2.4.2). Ensuite, une classification des gradients en trois classes (fort, moyen, faible), selon des critères statistiques (liés à la perception visuelle et à la nature des régions voisines) (cf. 3.3.1.1.1) permet d'associer un indice de confiance à chacun des gradients.

L'extraction des régions uniformes est basée sur l'approche d'agrégation de points selon une version améliorée de l'algorithme récursif présenté dans la section 3.2.2.1. Le principe consiste à regrouper les points de l'image suivant deux critères. Le premier tient compte du contraste local et de la variance de la région en cours de formation. Le second et nouveau critère fait intervenir l'indice de confiance associé au gradient du point examiné. Un point dont le niveau de gris est similaire à la région connexe (inférieur à un seuil réactualisé à chaque itération) et dont l'indice de confiance est faible, est agrégé à cette région.

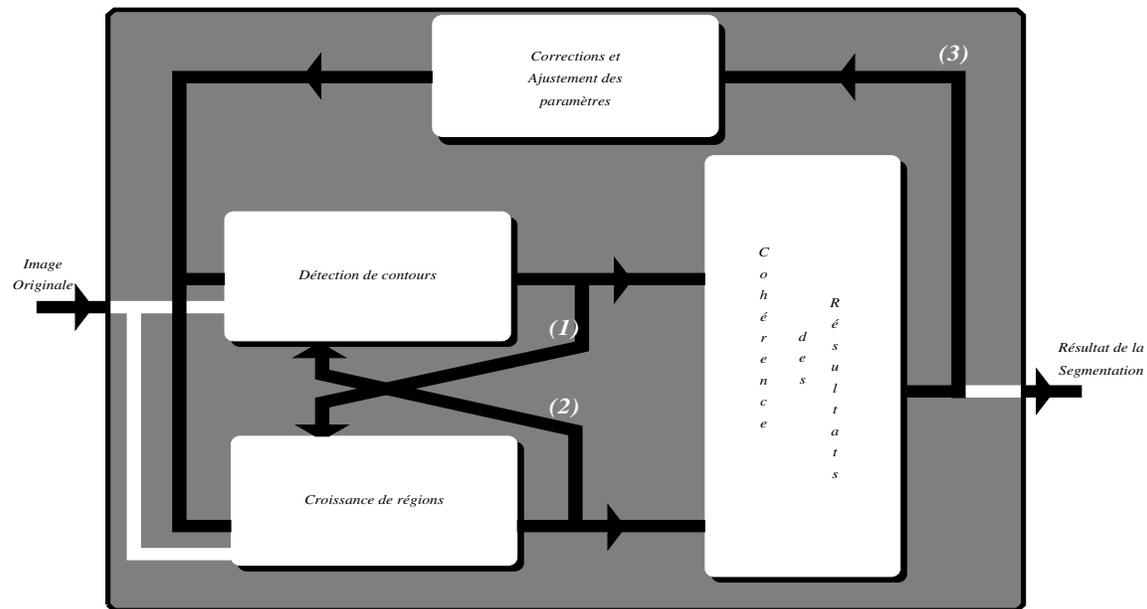


FIG. 3.7 - *Synoptique du système de segmentation : cas de la coopération contours-régions uniformes.*

La méthodologie proposée dans ce premier module du système tend à tirer un meilleur profit de la dualité entre les régions et les contours dans l'image. L'objectif est d'aboutir à deux résultats mutuellement compatibles afin d'obtenir une segmentation fiable. La dualité contours-régions uniformes peut être exprimée de quatre manières différentes [270] [365] [54] [184] :

- les régions sont situées à l'intérieur des contours, et par conséquent il n'y a pas de points de contours à l'intérieur d'une région.
- un point contour réel est situé sur ou à proximité d'une frontière de région.
- une frontière de région est fermée par nature, et un contour doit l'être également.
- un contour ne peut pas être situé à l'intérieur d'une région, et un contour doit être situé sur la totalité de la frontière commune entre deux régions.

Le processus de segmentation est itéré un certain nombre de fois (boucle de retour (3) de la figure 3.7). A la fin de chaque itération, la carte des régions est transformée en une carte de contours. Une mesure de similarité entre celle-ci et le résultat de la détection de contours est effectuée pour évaluer la cohérence des deux résultats. Cette mesure est réalisée à l'aide d'une métrique inspirée des travaux de Baddeley [18]. Le résultat de la segmentation qui sera retenu est celui qui minimise cette mesure.

L'algorithme général de segmentation dans sa version séquentielle peut être décrit comme suit [188] :

- Etape 1 :** *Initialisation par une sur-détection de contours; et une sur-segmentation de régions (noyaux)*
- Etape 2 :** *Croissance de régions autour des noyaux en utilisant le résultat de la détection de contours précédent;*
- Etape 3 :** *Corrections et affinements de la carte des contours  $C$  en utilisant le résultat de la croissance de régions;*
- Etape 4 :** *Transformation en contours du résultat de la croissance de régions pour obtenir une carte de frontières  $F$ ;*
- Etape 5 :** *Mesure du degré de dissimilarité  $Dissim_{cf}(C, F)$  entre la carte de frontières et la carte de contours; Si  $Dissim_{cf}(C, F) < DistMin$  alors  $\{DistMin := Dissim_{cf}(C, F)\}$ ; Résultat régions optimal:=  $F$ ; Résultat contours optimal:=  $C$ ; Si le critère d'optimalité est vérifié alors Fin; sinon continuer;*
- Etape 6 :** *Relâcher les contraintes; aller à **Etape 2**;*

Nous allons maintenant expliquer comment le processus de croissance de régions peut bénéficier de l'aide de la détection de contours et vice-versa. Nous évoquerons, ensuite, quelques techniques de mise en correspondance pour mesurer la corrélation entre plusieurs segmentations. Une présentation de la distance de Hausdorff et de celle de Baddeley ainsi que leurs propriétés sera effectuée. Nous présenterons, enfin, notre algorithme pour mesurer la similarité entre les résultats issus des deux méthodes de segmentations.

### 3.2.3.1 Interaction contours-régions uniformes

Le principe de la méthode consiste à itérer le partitionnement de l'image en commençant par des critères stricts et en effectuant des corrections des résultats intermédiaires. Puis les paramètres qui contrôlent les deux processus sont relâchés progressivement et la procédure globale est relancée jusqu'à la convergence vers un résultat commun (contours et régions). La correction des résultats intermédiaires et la coopération visent à tirer partie de la dualité entre les régions et les contours. Cette dualité est exprimée selon les règles vues précédemment. Sur la base de ces règles, des affinements des résultats intermédiaires sont effectués. Un certain nombre de critères décisionnels sont utilisés pour corriger les contours et régions obtenus à l'issue des différentes itérations.

#### 3.2.3.1.1 Croissance de régions avec contrôle des contours

La procédure de croissance de région admet en entrée, outre l'image originale pré-traitée, l'image des points de contours. Comme nous l'avons déjà vu dans la section 3.2.2.1, de nombreuses techniques existent pour détecter les régions uniformes dans une image. Connaissant

les faiblesses de chacune d'elles, il s'agit donc d'exploiter au mieux les informations sur les contours pour aider la croissance des régions dans les cas problématiques. Aussi, lors de la formation d'une région, si un pixel ne satisfait pas le critère d'agrégation, une vérification de l'existence effective de contours à proximité de la frontière de la région est effectuée. Dans le cas positif, les contours sont confirmés, et la croissance de la région est arrêtée (figure 3.8).

FIG. 3.8 - *Croissance de région autour de noyaux homogènes guidée par les contours.*

### 3.2.3.1.2 Affinement et correction des contours en utilisant les régions

La difficulté principale de la détection de contours est liée à la présence des points bruités et des contours non significatifs. Ce type de contours reflète entre autres des défauts relatifs à l'état de surface des objets (textures) ou à la présence d'ombres et de reflets. La présence du bruit contribue à la production des faux contours et engendre des irrégularités locales des contours.

Pour corriger ou réduire ces effets indésirables, deux types de traitements sont généralement nécessaires :

- un pré-traitement par lissage de l'image originale afin de régulariser les données [332];
- un post-traitement par analyse de configuration locales des contours.

Dans notre cas, les points candidats aux contours retenus sont ceux correspondant aux maxima locaux de l'image des gradients. Même si ces maxima comprennent des points de faux contours, ils présentent l'avantage de préserver des informations nécessaires pour aider à la fermeture des contours et pour donner une indication précise sur leur position.

L'extraction des maxima locaux est effectuée sur un voisinage de taille  $3 \times 3$  de l'image norme du gradient. Le but est d'extraire tous les pixels présentant un maximum local dans la direction du gradient. Afin d'éviter que cette étape ne soit trop pénalisante du point de vue du temps de calcul, l'information orientation n'est en fait calculée qu'à partir du rapport des gradients directionnels  $g_x(i, j)$  et  $g_y(i, j)$ . Cette étape est illustrée par la figure 3.9 qui

représente un voisinage de taille 3 x 3 dans l'image norme du gradient.

FIG. 3.9 - *Extraction des maxima locaux.*

Pour tout point de l'image des gradients  $\sqrt{g_x^2(i, j) + g_y^2(i, j)}$  faire :

On définit  $H$  par :

$$H = \frac{g_x(i, j)}{g_y(i, j)} \quad (3.8)$$

Si  $g_y(i, j) > 0$ , alors le point central, dont la norme du gradient est notée  $G_R$  dans la figure 3.9, sera déclaré comme maximum local si

$$G_R > U \quad \text{et} \quad G_R \geq Z \quad (3.9)$$

Les variables  $U$  et  $Z$  données par :

$$\begin{aligned} \text{Si } g_x(i, j) < 0 \quad \text{et} \quad g_y(i, j) = 0 \quad \text{alors} \quad U = P_8, \quad Z = P_4 \\ \text{Si } g_x(i, j) > 0 \quad \text{et} \quad g_y(i, j) = 0 \quad \text{alors} \quad U = P_4, \quad Z = P_8 \end{aligned} \quad (3.10)$$

$$\begin{aligned} \text{Si } H > 1 \quad U &= ((H - 1)P_4 + P_5)/H, \quad Z = ((H - 1)P_8 + P_1)/H \\ \text{Si } 1 > H \geq 0 \quad U &= (1 - H)P_6 + HP_5, \quad Z = (1 - H)P_2 + HP_1 \\ \text{Si } 0 > H \geq -1 \quad U &= (1 + H)P_6 - HP_7, \quad Z = (1 + H)P_2 - HP_3 \\ \text{Si } -1 < H \quad U &= ((H + 1)P_8 - P_7)/H, \quad Z = ((H + 1)P_4 - P_3)/H \end{aligned} \quad (3.11)$$

Si  $g_y(i, j) < 0$ , la même procédure est à effectuer mais en permutant  $U$  et  $Z$ .

Après l'extraction de maxima locaux, une classification des gradients en trois catégories (contraste faible, moyen et fort), selon des critères liés à la perception visuelle (cf. 3.3.1.1.1), permet d'associer un indice de confiance à chacun des gradients [83].

Pour corriger les irrégularités des contours, une analyse des configurations locales des contours est effectuée. Cette opération est réalisée selon le procédé suivant : si dans la carte des points de contours candidats, le point examiné (point en noir Fig.3.10(a)) et ses voisins

se trouvent, entre autres, dans l'une des situations de la figure 3.10 (a) alors, il est déplacé ou bien supprimé selon le cas, pour obtenir la configuration de la figure 3.10 (b). Ce traitement est généralisé en considérant toutes les autres configurations locales obtenues par rotation de celles présentées par la figure 3.10(a).

FIG. 3.10 - *Correction des irrégularités locales.*

Le bruit provoque également de faux contours autour d'un vrai contour, notamment dans le cas des points de contours situés dans un environnement peu texturé. Pour remédier aux problèmes engendrés par ce type de contours et par les autres contours non significatifs, un traitement spécifique est effectué au cours du processus coopératif de segmentation après analyse des régions voisines qui les bordent. Les contours bordés par des régions de nature identique sont candidats à être éliminés.

Le processus de détection de contours exploite aussi les informations sur les régions pour confirmer et corriger les contours extraits. La recherche d'informations sur les régions voisines est effectuée sur des fenêtres d'analyse locales. Les informations étant généralement plus globales, les contours peuvent être affinés et la détection devient alors plus fiable. Un certain nombre de critères sont employés pour atteindre cet objectif :

- si les régions situées de part et d'autre du pixel-contour examiné possèdent des labels différents, le point contour est validé.
- sinon, on examine le gradient du point de contour ainsi que ses voisins. Si les points de contours ont des gradients faibles, le contour est systématiquement rejeté.
- si le gradient appartient à la classe des gradients à fort contraste, le contour peut être facilement détecté sans avoir recours aux informations sur les régions voisines. En revanche, s'il appartient à la classe des gradients "moyens", les informations sur les régions voisines sont requises pour le valider ou le rejeter.

L'autre aspect de l'interaction contours-régions est celui lié à la fermeture de contours. Celle-ci consiste à combler les discontinuités entre des suites de points représentant le même contour. Ces discontinuités sont causées par :

- une faible variation des luminances entre régions (gradient trop faible),

- un phénomène purement photométrique (reflet, ombre, réflexion, ...) occultant le contour,
- les déformations optiques ou le bruit électronique du système d'acquisition (objectif, CCD, numérisation, ...).

La fermeture de contours est généralement initialisée par un regroupement perceptuel. Une première étape consiste à examiner les chaînes de contours à l'aide des critères de Gestalt [216]. Une fois la fermeture décidée, chaque paire de chaînes de points de contours concernée est fusionnée par une courbe, définie avec ou sans critère photométrique caractérisant le trou séparant ces chaînes.

L'objectif recherché est la description de l'image par association de primitives liées entre elles par des propriétés géométriques et/ou topologiques [216]. Cette approche met en oeuvre des mécanismes proches de ceux de la vision humaine. En effet, l'idée de base consiste à remarquer que, dans une image, l'oeil est capable de détecter facilement les segments de droite, les arcs de cercle ainsi que les groupements de segments colinéaires, parallèles, connexes ou symétriques noyés dans un ensemble de segments disposés aléatoirement (figure 3.11).

FIG. 3.11 - *Groupement perceptuel des contours.*

Généralement, les critères de regroupement de Gestalt sont donnés par :

- la proximité : deux chaînes ne peuvent être fusionnées si leur distance est supérieure à un seuil;
- la continuité : le contour étant souvent rectiligne, deux chaînes ont une forte probabilité de se fusionner si elles sont dans le prolongement l'une de l'autre.

La courbe reliant deux chaînes de points de contours est définie soit en utilisant un modèle de la courbe (arc de cercle, polynômes, parabole, segment de droite, ...), soit en minimisant une fonction d'énergie de courbure. La fermeture est fondée sur des critères purement locaux aux deux chaînes concernées [119] [249] ou sur une stratégie globale d'optimisation [13] [97] [331] [372].

Notre implantation actuelle de la fermeture est simple. Deux chaînes de points de contours sont reliées si :

- la distance qui sépare deux de leurs extrémités est faible (inférieure à un seuil),
- au moins l'une des deux chaînes de points de contours concernées se situe sur ou à proximité d'une frontière de région.

Différents cas de figures se présentent. En cas de jonction simple de type L ou V, nous relierons leurs extrémités par un segment de points de contours. Toutefois, cette approche échoue dans le cas de jonction multiple de type Y ou X. Il est bien sûr possible d'effectuer une recherche sur des jonctions de cette nature, en calculant le barycentre des extrémités et en reliant ce barycentre à chaque extrémité par un segment. Toutefois, cette opération nécessite un coût calculatoire élevé.

Si un contour n'est pas fermé, ou s'il est inclus dans une même région, celui-ci est conservé s'il appartient à la classe des gradients à contraste fort. Il peut donc arriver qu'un contour reste présent au milieu d'une région. Ce cas de figure est rencontré dans les régions texturées à motifs fortement contrastés. Ce problème sera, en partie, traité dans la section 3.3.

### 3.2.3.2 Mesure de la correspondance contours-régions

#### 3.2.3.2.1 Position du problème

Etant données deux segmentations d'une même scène, le problème est de trouver une mesure qui permet de donner une estimation de la similarité ou de la correspondance entre les éléments des deux segmentations, en considérant à la fois la détection et la topologie des différentes primitives extraites.

Avant de présenter notre approche pour mesurer cette correspondance, nous donnons dans les paragraphes suivants quelques exemples de méthodes de mise en correspondance faisant coopérer différents traitements. Pour une revue plus exhaustive, on peut se reporter à [33] et plus récemment à [102] et [143].

Dans [151], Hoff et Ahuja détectent des contours par passage par zéro, pour chacun des niveaux d'une pyramide de résolution, des résultats de l'application d'un opérateur de type laplacien de gaussienne de taille variable. Ils soutiennent l'idée que les étapes de mise en correspondance et d'interpolation puissent être traitées simultanément. Dans la recherche des correspondances, ils considèrent que les solutions candidates doivent appartenir à une surface plane. Cette surface, déterminée au niveau précédent, permet de trouver les "bons" correspondants (utilisation de la transformée de Hough). Ceux-ci sont alors utilisés pour déterminer un modèle de facette (surface quadratique) plus précis, au niveau courant. Partant de ce modèle, ils cherchent ses discontinuités d'orientation, pour trouver les contours d'occultations sous forme de splines cubiques.

Dans [244], les auteurs décrivent la courbe issue d'une ligne de signal sous forme de plateaux, pics, vallées. Une fonction d'évaluation pour chacun des trois niveaux de résolution de la pyramide est créée, liée à la forme de la ligne épipolaire, à sa description symbolique et à son intensité. Une relation d'intra/inter-scanline permet d'imposer une certaine consistance de la disparité entre les lignes épipolaires. La courbe d'intensité précédemment utilisée est ensuite convertie en courbe de gradient et la même technique est à nouveau utilisée à partir du résultat obtenu sur la courbe d'intensité. Un post-traitement est alors activé pour le problème des occultations. Enfin, une interpolation permet de compléter la carte.

L'utilisation de codages de chaînes pour la recherche de formes de primitives est à la base de l'approche décrite dans [3]. Après extraction des zones non texturées dans les deux images, un code chaîne puis un code chaîne de fréquence sont successivement appliqués sur l'ensemble de leurs contours classés par taille. Les correspondants sont alors recherchés trois par trois dans la seconde image. Cette recherche de triplés apporte une notion de géométrie relative entre les formes. Différents critères d'orientation, de surface et de longueur, permettent aux auteurs d'identifier le bon correspondant de chaque triplé. Une fois le triplé correspondant trouvé, l'itération suivante crée un nouveau triplé à partir de deux zones déjà appariées dans le triplé précédent et d'une nouvelle non appariée. Pour McKeown et Hsieh [244], le codage ne s'applique pas aux contours, mais à l'ensemble des lignes des images gauche et droite (après correction en géométrie épipolaire).

Dans [169], une méthode de mise en correspondance de segments, utilisant un graphe décrivant les relations structurelles de ces segments entre eux, est proposée. Chaque segment est orienté en fonction du contraste à sa droite et à sa gauche. Les segments forment alors les nœuds du graphe relationnel, liés à leur agencement dans l'image. La mise en correspondance des deux graphes image gauche et droite, se fait alors par la recherche de cliques maximales dans un nouveau graphe décrivant des contraintes géométriques.

A partir de segments extraits sur la paire stéréo rectifiée, la méthode de [236] cherche pour un segment donné, à minimiser la différence de disparité des segments appartenant à une fenêtre centrée sur celui-ci. Cette optimisation assimilable à une relaxation, est exprimée par l'intermédiaire d'une fonction convergent par itérations successives, vers un état de pseudo stabilité. En pratique, un nombre maximal d'itérations est fixé.

Dans [295], les images de la séquence ne sont pas rectifiées contrairement à [88]. Un détecteur de "coins" sur l'image courante associe dans l'image suivante une série de coins pour la bande épipolaire associée (erreur de rectification, géométrie épipolaire). La bande de recherche de disparité étant limitée (disparité min et max), il en est de même pour le nombre de coins associés. L'orientation et le type du coin permettent alors d'éliminer un maximum de candidats. Un graphe de relation entre les coins (présence ou non de segment) d'une même image est créé avec : (i) un lien "fort" si l'azimut et le gradient d'intensité du segment sont relativement constants dans le couple d'images, (ii) un lien "faible" si le

gradient varie, (iii) pas de lien dans les autres cas. Les nœuds du graphe étant constitués par les couples de coins mis en correspondance, la recherche des surfaces est alors équivalente à une recherche de cycle dans le graphe. Pour limiter les recherches, seuls les coins 3D de valeur d'angle de 90 degrés sont conservés. Les segments 3D (lien fort) parallèles ou formant un angle droit servent alors d'origine à la recherche de cycles d'un maximum de huit coins coplanaires. Une contrainte d'homogénéité des surfaces permet de conserver celles qui sont les plus vraisemblables.

Dans les techniques de la programmation dynamique discrète, on ne pouvait choisir qu'un chemin passant par les nœuds du graphe [253] et les résultats en étaient affectés. Dernièrement, une version continue a été présentée dans [307]. Elle se base sur un critère de moindre déformation géométrique, à l'exclusion de toutes autres contraintes telles que: l'utilisation de lignes épipolaires ou la paramétrisation de la déformation en analyse du mouvement.

Dans les techniques de vote [220], on définit un score de similarité, fonction de caractères géométriques et topologiques des chaînes de contours. Le correspondant est validé s'il obtient le meilleur score dans les deux sens gauche vers droite et droite vers gauche, une structure pyramidale lui permettant une amélioration des résultats.

Dans notre cas, nous nous intéressons à la mesure du degré de similarité entre résultats de segmentations. Pour des raisons de simplicité et de temps de calcul, une représentation sous forme de primitives de type contours a été adoptée. Notons qu'une modélisation de plus haut niveau (basée sur l'utilisation de primitives géométriques de types segments de droite ou autres) serait beaucoup plus fiable et plus significative et permettrait d'obtenir de meilleurs résultats dans un environnement peu texturé et présentant de fortes discontinuités. Elle nécessiterait, cependant, l'utilisation de techniques d'interpolation et de méthodes de mesures de similarité plus complexes, ce qui engendrerait un coût algorithmique beaucoup plus important.

### 3.2.3.2.2 Mesure de similarité entre contours et régions

Les deux paragraphes suivants présentent quelques distances utilisées pour mesurer la similarité entre images de contours. Nous exposons ensuite l'approche que nous avons développée pour la mesure de la dissimilarité entre cartes de contours.

#### • Distance de Hausdorff

La distance de Hausdorff entre deux sous ensembles  $A, B$  appartenant à un ensemble de points  $X$  est définie par :

$$D_H(A, B) = \max\left\{\sup_{x \in A} d_E(x, B), \sup_{x \in B} d_E(x, A)\right\} \quad (3.12)$$

où si  $d$  est une distance entre deux pixels alors la distance  $d_E(x, A)$  entre un pixel  $x$  et un ensemble de pixels  $A$  est définie par [86] :

$$d_E(x, A) = \inf\{d(x, a) : a \in A\} \quad (3.13)$$

Bien que la distance de Hausdorff soit théoriquement intéressante puisqu'elle possède des propriétés liées avec les opérations de base de la morphologie mathématique [161], elle est très peu utilisée en pratique comme mesure d'erreur entre images binaires.

Le problème inhérent à cette distance est surtout dû à sa grande sensibilité aux bruits (de contours). Un simple pixel mal classé augmente la valeur de  $D_H$  à cause de la présence de l'opérateur "sup" dans sa définition. L'exemple de la figure 3.12 illustre cette sensibilité.

FIG. 3.12 - *Sensibilité de la distance de Hausdorff au bruit.*

Une manière de réduire cette sensibilité de  $D_H$  aux grandes valeurs d'erreur causées par quelques pixels est de modérer les valeurs de  $d$  entre deux pixels par l'application d'une transformation  $w(x)$  sur les valeurs de  $d$  de façon à ce qu'elle préserve les propriétés d'une distance [18].

Comme exemple de  $w(x)$  on peut avoir :  $\frac{c \cdot x}{x+1}$ , ou encore  $\min(x, c)$  avec  $c > 0$ , ...

Ces fonctions ont pour effet de limiter les valeurs de distances à une valeur limite  $c$ .

L'application de la transformation  $w$  sur la distance entre un pixel et un sous ensemble génère la distance  $d_w$  appelée "distance contrainte".

Si  $d$  est une distance entre pixels et  $d_E(x, A)$  la distance entre le pixel  $x$  et l'ensemble  $A$ , induite par  $d$ , on a :

$$d_w(x, A) = w(d_E(x, A)) \quad (3.14)$$

Si  $H_d$  est la distance de Hausdorff obtenue avec la distance  $d$  alors :

$$H_w(A, B) = w(H_d(A, B)) \quad (3.15)$$

### • Distance de Baddeley

Comme le montre schématiquement l'exemple de la figure 3.12, selon la distance de Hausdorff, les deux formes  $A$  et  $B$  ayant une forte similitude, se voient affecter la distance maximale  $D_H(A, B)$  qui sépare les points les plus éloignés.

Afin d'éviter cette sensibilité de  $D_H$  au bruit et compte tenu de ses propriétés topologiques intéressantes, Baddeley propose dans [18] de remplacer la distance du sup par une simple moyenne ou une moyenne d'ordre  $q$ . Il montre tout d'abord que la distance de Hausdorff peut s'écrire sous la forme :

$$D_H(A, B) = \sup_{x \in X} |d_E(x, A) - d_E(x, B)| \quad (3.16)$$

En remplaçant le "sup" par une moyenne d'ordre  $q$ , il obtient une nouvelle distance appelée  $D^q$  :

$$D^q(A, B) = \left[ \frac{1}{N} \sum_{x \in X} |d_E(x, A) - d_E(x, B)|^q \right]^{1/q} \quad (3.17)$$

où  $N = \text{card}(X)$  est le nombre de pixels dans l'image, et  $1 \leq q < \infty$ .

En modifiant la distance entre deux pixels par la fonction  $w(x)$  continue, concave, strictement croissante, et telle que  $w(0) = 0$ , la distance de Baddeley s'écrit alors :

$$D_w^q(A, B) = \left[ \frac{1}{N} \sum_{x \in X} |w(d_E(x, A)) - w(d_E(x, B))|^q \right]^{1/q} \quad (3.18)$$

La fonction  $w(x) = \frac{c \cdot x}{x+1}$  a pour effet d'accorder une contribution négligeable aux pixels éloignés de  $A$  et  $B$ . La distance de Baddeley ne dépend alors que du voisinage immédiat de  $A$  et  $B$ . Dans le cas de la fonction  $w(x) = \min(x, c)$ , on ramène à zéro la contribution des pixels éloignés de  $A$  et  $B$  d'une distance supérieure à  $c$ .

Les paramètres  $c$  et  $q$  déterminent un compromis entre les erreurs de détection et les erreurs de localisation.

Pour des valeurs faibles de  $c$ , la distance de Baddeley se comporte comme une mesure du pourcentage des pixels mal classés.

La valeur de  $q$  détermine l'importance relative donnée aux valeurs élevées de l'erreur de localisation. Pour  $q \rightarrow \infty$ , la distance de Baddeley devient une distance de Hausdorff :

$$D_w^\infty(A, B) = H_w(A, B) \quad (3.19)$$

- **Mesure de dissimilarité développée**

Nous décrivons dans ce qui suit la procédure utilisée pour mesurer la dissimilarité entre contours et régions. Soit  $C$  la carte des contours et soit  $F$  celle des frontières des régions. L'approche adoptée consiste à calculer pour chaque point de contour dans l'image  $C$  (resp.  $F$ ) l'écart minimal qui le sépare d'un point frontière dans l'image  $F$  (resp.  $C$ ). On calcule, ensuite, les moyennes des écarts minimaux des deux cartes. La dissimilarité finale est donnée par la moyenne des deux moyennes.

La mesure de la similarité entre la carte des contours notée  $C$  et celle des frontières des régions  $F$  est effectuée selon l'algorithme suivant :

**Pour tout point contour  $c$  de  $C$**

**Se positionner au point  $c'$  de  $F$  ayant les mêmes coordonnées de  $c$ .**

**Rechercher dans  $F$  le point contour  $v'$  qui minimise une distance  $d_e(c', v')$  telle que si  $c'$  et  $v'$  ont pour coordonnées respectives  $(i, j)$  et  $(k, l)$  :**

$$d_e(c', v') = \sqrt{(i - k)^2 + (j - l)^2}$$

**On note  $d_{min}(c, F)$  l'écart minimal.**

**Pour tous les points contour  $f$  de  $F$ , les mêmes opérations sont répétées :**

**Projeter le point  $f$  sur la carte  $C$ . Soit  $f'$  ce point.**

**Rechercher dans  $C$  le point contour le plus proche du point  $f'$  au sens de la distance euclidienne  $d_e(., .)$ . On note  $d_{min}(f, C)$  cet écart minimal.**

**La dissimilarité entre les deux cartes  $C$  et  $F$  est donnée par**

$$Dissim(C, F) \equiv 1/2\{\langle d_{min}(c, F) \rangle_C + \langle d_{min}(f, C) \rangle_F\}$$

où  $\langle d_{min}(c, F) \rangle_C$  correspond à la moyenne des écarts  $d_{min}(c, F)$

et  $\langle d_{min}(f, C) \rangle_F$  correspond à la moyenne des écarts  $d_{min}(f, C)$

$$\langle d_{min}(c, F) \rangle_C = 1/N_c \sum_{c \in C} d_{min}(c, F) P(d_{min}(c, F))$$

$$\langle d_{min}(f, C) \rangle_F = 1/N_f \sum_{f \in F} d_{min}(f, C) P(d_{min}(f, C))$$

$P(x)$  désigne la population de  $x$ .

$N_c$  (resp.  $N_f$ ) correspond au nombre de points de contours dans  $C$  (resp.  $F$ ).

Cet algorithme donne de bons résultats : il présente une forte robustesse au bruit comparativement à la distance de Hausdorff, et ne nécessite aucun seuil empirique. Il reste néanmoins très coûteux en temps de calcul. Pour contourner en partie ce problème, nous avons réduit l'espace de recherche des points les plus proches de  $p$  dans  $M$  ( $p = c'$  ou  $f'$  et  $M = F$  ou  $C$ ) à des voisinages limités et centrés en  $p$  comme illustré dans la figure 3.13.

FIG. 3.13 - *Espace de recherche du point le plus proche.*

### 3.2.4 Résultats expérimentaux

Afin de valider le premier module, nous ne nous sommes pas fixés un type d'images précis. Nous avons considéré des images provenant de diverses banques d'images et en particulier celle du GdR-PRC ISIS. Ces images sont de taille 256x256 ou 512x512 pixels et sont toutes codées sur 256 niveaux de gris. La non introduction d'information a priori était une contrainte fixée. L'algorithme est entièrement non-supervisé, et ne dépend pas du type d'image choisi. En effet, les résultats sont obtenus aussi bien sur des images réelles que sur des images synthétiques, sans modifier les paramètres de la méthode.

Avant de présenter les résultats obtenus sur quelques images de scènes réelles, nous présentons dans la figure 3.14 (resp. 3.15) l'évolution des segmentations contours et régions du quart supérieur gauche de l'image Aquitaine (image Quart-Aquitaine) lors des premières itérations (resp. dernières itérations). La première ligne de la figure 3.14 présente respectivement l'image originale, l'image pré-traitée, l'image des gradients et l'image des contours du départ. Les lignes suivantes représentent les différents résultats intermédiaires obtenus à l'issue des différentes itérations. La première colonne présente l'évolution des résultats de la croissance des régions, les images de la deuxième et troisième colonne correspondent respectivement aux images contours et frontières de régions. La dernière colonne présente la superposition des contours et des frontières. Nous pouvons remarquer la formation des régions noyaux homogènes lors des premières itérations, les points de l'image étant ensuite progressivement agrégés autour de ces différents noyaux jusqu'à ce que les frontières atteignent les contours validés avoisinants.

Dans cette expérimentation ainsi que dans l'ensemble des expérimentations suivantes, le paramètre de croissance de région  $S_k$  débute à la valeur 1 et croît par pas de 1 à chaque itération. Le paramètre de surface  $N_{R_k}$  initialisé à 32 décroît avec un pas de 1. Le facteur d'échelle de la détection de contours  $\alpha$  a été fixé à 2.0. La figure 3.16 montre l'évolution de la mesure de similarité normalisée (par rapport à la distance maximale) entre la carte des contours et la carte des frontières des régions.

FIG. 3.14 - *Segmentations en contours et en régions intermédiaires, image Quart-Aquitaine.*

FIG. 3.15 - *Suite, segmentations intermédiaires et finales en contours et en régions, image Quart-Aquitaine.*

No. d'itération	1	2	3	4	5	6	7	8
D(C,F)	0.183663	0.012594	0.006653	0.004944	0.004569	0.004377	0.004344	<b>0.004249</b>

No. d'itération	9	10	11	12	13	14	15	16	17
D(C,F)	0.004325	0.004310	0.004290	0.004256	0.004271	0.004256	0.004365	0.004400	0.004399

FIG. 3.16 - *Evolution de la dissimilarité entre la carte des contours et la carte des frontières de régions en fonction des itérations. Image Quart-Aquitaine.*

Les figures 3.17 3.18 et 3.19 représentent respectivement les résultats de segmentation de l'image "Aquitaine" complète et l'évolution de la dissimilarité. Cette évolution montre que le résultat de segmentation commence à se stabiliser à partir de la 6ième itération. Nous pouvons remarquer que pour cette image, de bons résultats de segmentation sont obtenus. La plupart des régions de l'image originale visibles à l'oeil sont détectées de manière satisfaisante. Les contours fins correspondent bien aux discontinuités visibles dans l'image originale, et la majorité d'entre eux sont extraits. On peut également noter leur bonne localisation et la préservation des structures fines.

(a) (b)

FIG. 3.17 - *Segmentation coopérative contours-régions. (a) Image Aquitaine. (b) Résultat final de la segmentation région.*

(a) (b)

FIG. 3.18 - *Segmentation coopérative contours-régions. (a) Carte de contours de départ. (b) Frontières finales de la segmentation région.*

No. d'itération	1	2	3	4	5	6	7	8
D(C,F)	0.135410	0.007438	0.003302	0.002556	0.002256	0.002105	0.002077	0.002050

No. d'itération	9	10	11	12	13	14	15	16	17
D(C,F)	<b>0.002024</b>	0.002056	0.002057	0.002048	0.002085	0.002082	0.002144	0.002151	0.002151

FIG. 3.19 - *Evolution de la dissimilarité entre la carte des contours et la carte des frontières de régions en fonction des itérations. Image Aquitaine.*

Nous présentons dans les figures qui suivent des résultats obtenus sur d'autres images réelles issues de différentes scènes.

FIG. 3.20 - *Segmentation coopérative contours-régions. Image IRM.*

FIG. 3.21 - *Segmentation coopérative contours-régions. Image Car.*

FIG. 3.22 - *Segmentation coopérative contours-régions. Image Aérienne.*

Nous pouvons constater que pour l'image de scène d'extérieur (figure 3.21) et l'image IRM (figure 3.20), les frontières des régions sont bien détectées et permettent donc d'envisager un processus d'interprétation de manière quasi-directe. Cependant, l'image aérienne (figure 3.22) segmentée présente une multitude de petites régions difficilement exploitables. La méthode de segmentation nécessite l'intégration d'autres types d'informations basées sur la texture, d'où l'intérêt de l'introduction du deuxième module dans le système développé.

### 3.3 Intégration de l'information texture dans la coopération contours-régions

Une amélioration, permettant de prendre en compte l'aspect textuel, est apportée en intégrant un module pour la localisation des primitives de type "textures". Celui-ci permet de corriger les primitives contours non significatives issues de la coopération contours-régions uniformes. La construction de la carte des primitives "textures" s'effectue à l'aide d'un algorithme de segmentation de textures opérant suivant deux niveaux : division et fusion. La fusion s'appuie sur les co-occurrences de niveaux de gris. Celle-ci présente l'inconvénient, dans son implantation classique, de conduire à des temps de calcul élevés. Pour pallier ce problème, le nombre de niveaux de gris initial est réduit par une étape de division originale. Cette étape est réalisée à l'aide d'une technique de multi-seuillage utilisant des critères d'homogénéité locale prenant en compte des propriétés du système visuel humain liées à la sensibilité de l'oeil aux contrastes. Le synoptique général du système dans ce cas est donné par la figure 3.23.

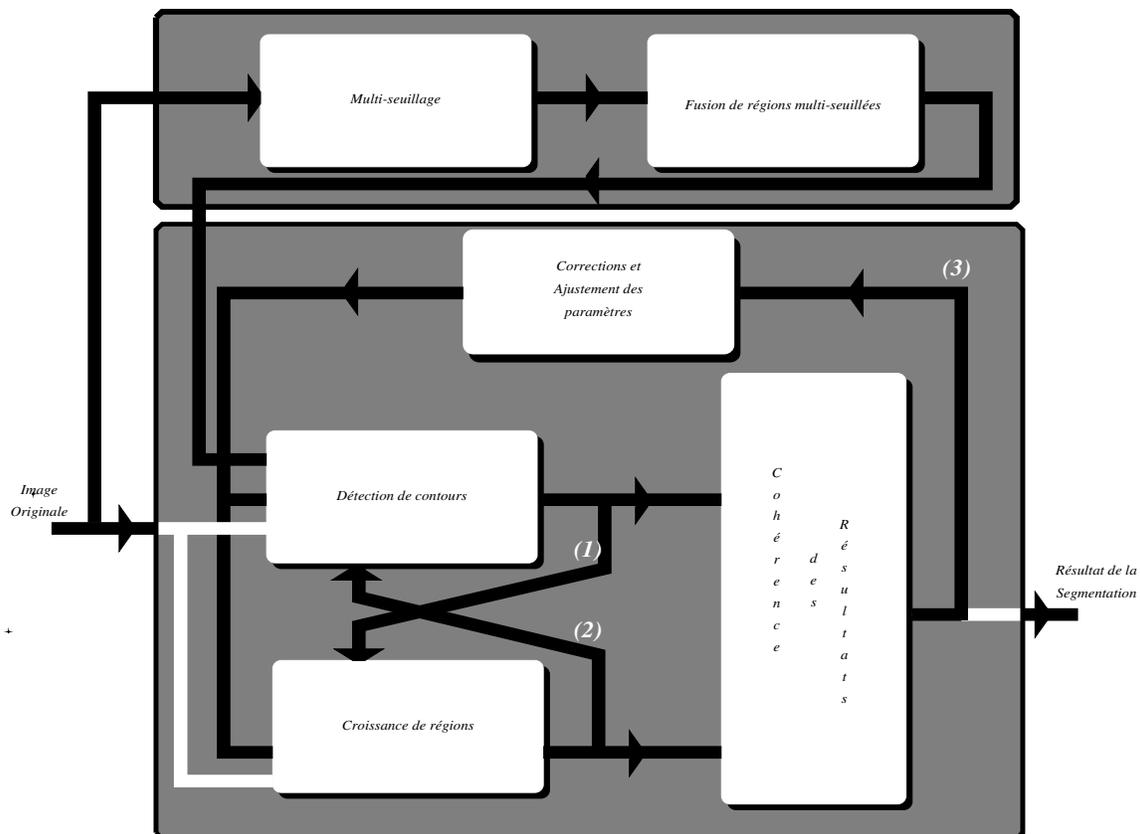


FIG. 3.23 - Synoptique du système de segmentation : cas de la coopération Contours-Régions-Textures.

Avant de décrire l'interaction entre les deux modules, nous présentons la méthode de localisation de textures développée et quelques résultats d'évaluation.

### 3.3.1 Localisation des régions texturées

Afin de localiser les primitives de type "textures" dans l'image, un algorithme de segmentation fondé sur une approche à deux niveaux (division et fusion) a été développé.

L'objectif de la *division* est de fournir une représentation cohérente et simplifiée au sens des niveaux de gris et suffisamment représentative pour la caractérisation des textures de l'image. Elle est réalisée à l'aide d'une technique de multi-seuillage. Les résultats de cet algorithme vont ainsi constituer les données d'entrée du processus de fusion.

La *fusion* a pour objectif de détecter de manière grossière les régions texturées de l'image. Les différentes études réalisées jusqu'à ce jour sur des images de scènes réelles ont montré que la plupart des méthodes d'analyse de textures sont insensibles aux relations spatiales supérieures au second ordre<sup>1</sup>. En conséquence, les paramètres du premier et second ordre sont souvent suffisants pour caractériser une texture. Aussi dans notre cas, la procédure de fusion est basée sur les co-occurrences de niveaux de gris issus de la division.

Dans les paragraphes suivants nous décrivons avec plus de détails ces deux procédures.

#### 3.3.1.1 Simplification par multi-seuillage

L'objectif de cette méthode est de fournir une représentation simplifiée tout en préservant les primitives pertinentes au sens des niveaux de gris. Le résultat produit sera utilisé par la procédure de localisation de textures présentée dans la section 3.3.1.2.

L'approche proposée consiste à classer les points de l'image en comparant leurs niveaux de gris aux seuils déterminés par l'analyse d'un histogramme global. Celui-ci est construit à partir de pics significatifs issus d'histogrammes locaux. Cette procédure élimine l'information non pertinente afin d'homogénéiser les régions et de mettre en évidence les éléments les plus caractéristiques des textures. Ceci permettra ultérieurement une simplification des procédures de segmentation ou de modélisation.

La procédure est effectuée en deux phases: *construction de l'histogramme global des maxima locaux* et *classification des niveaux de gris de l'image à partir de l'histogramme global*.

---

1. On note toutefois que dans une étude consacrée à la discrimination de textures, Julesz a donné des exemples qui montrent que l'oeil est incapable de discerner certaines textures artificielles différentes ayant des caractéristiques de premier et second ordre identiques [179].

- *La première phase* s'appuie sur la transformation des histogrammes locaux et sur des critères de décision simples pour valider les différents maxima locaux. La procédure de transformation est nécessaire car la plupart des histogrammes locaux multimodaux présentent des modes aux limites mal définies. Le calcul des maxima est alors délicat. Cette transformation a donc pour but de mettre en évidence les différents modes de chaque histogramme. Pour cela nous avons utilisé quatre critères utilisant une fonction de sensibilité de l'oeil au contraste. Ceux-ci donnent une indication sur l'homogénéité d'un point par rapport à ses voisins les plus proches. En fonction du nombre de critères vérifiés, les fréquences d'occurrence d'un histogramme donné sont transformées. Pour valider ou rejeter les maxima d'un histogramme, nous avons introduit des critères adaptatifs qui tiennent compte des distances entre maxima et de leurs amplitudes. Les différents maxima validés sont cumulés dans l'histogramme global.

La transformation des histogrammes et la détection des maxima locaux sont effectuées sur des fenêtres de différentes tailles, ce qui permet de tenir compte de la résolution de chaque image et de rendre le traitement adaptatif.

- L'objectif de *la deuxième phase* est de détecter les seuils optimaux de segmentation. Elle est réalisée moyennant une procédure itérative de regroupement des masses de l'histogramme global en masses ponctuelles significatives caractérisant les régions qui composent l'image. Les valeurs des seuils alors déterminées entre ces masses permettent de définir les intervalles de luminance de chaque classe et ainsi d'associer à chaque point de l'image une certaine étiquette.

La technique de multi-seuillage tient donc compte des propriétés locales de chaque histogramme et ne demande aucune connaissance a priori sur l'image à segmenter.

Dans ce qui suit nous décrivons les différentes étapes de cet algorithme.

### 3.3.1.1.1 Construction de l'histogramme global

Les histogrammes locaux sont transformés de manière à mettre en évidence leurs principaux modes. Les modes validés serviront alors pour la construction de l'histogramme global des maxima locaux. Cette série d'opérations est effectuée en utilisant plusieurs tailles de fenêtres (16x16, 32x32, 64x64) afin de tenir compte de la résolution spatiale de chaque image et rendre ainsi le traitement adaptatif.

La procédure de la construction de l'histogramme global est composée des deux étapes suivantes: *la transformation des histogrammes locaux et la recherche des pics significatifs.*

• **Transformation des histogrammes locaux**

L'image originale notée  $I$ , de taille  $N \times N$  est subdivisée en fenêtres rectangulaires sans chevauchement  $B_c$ , chacune de taille  $d \times d$ . Soit  $h_c(g_i)$  la fréquence d'occurrence du niveau de gris  $i$  dans la fenêtre  $B_c$ .

Pour accentuer de manière significative la séparation des modes les plus importants de l'histogramme, une procédure de transformation est employée. Cette procédure utilise quatre critères basés sur la mesure des fluctuations des niveaux de gris. Ces critères notés  $C_1, C_2, C_3$  et  $C_4$ , donnent une indication sur l'homogénéité d'un point par rapport à ses quatre voisins les plus proches. Ils sont respectivement définis par

$$|I(x, y) - I(x - 1, y)| < \Delta[I(x, y)]$$

$$|I(x, y) - I(x, y - 1)| < \Delta[I(x, y)]$$

$$|I(x, y) - I(x, y + 1)| < \Delta[I(x, y)]$$

$$|I(x, y) - I(x + 1, y)| < \Delta[I(x, y)]$$

où  $I(x, y)$  représente le niveau de gris du point  $(x, y)$  appartenant à  $B_c$ .

$\Delta[I(x, y)]$  est une fonction liée à la perception visuelle. Elle est obtenue expérimentalement et permet de décrire la sensibilité de l'oeil au contraste (fonction SOC) dans différents intervalles de niveaux de gris. La mesure de la sensibilité de l'oeil au contraste a été réalisée selon le procédé suivant [210]. Une série d'images composées de bandes verticales est créée (Fig.3.24). Chaque bande dans l'image correspond à un niveau de gris constant. Soit  $n = 0$  la valeur du niveau de gris de la première bande. Le niveau de gris de la seconde bande est obtenue en augmentant le niveau de gris de la première d'une valeur  $C$ ; celui de la troisième bande est de  $n + 2C$  et ainsi de suite. Le niveau de gris de la dernière bande est égale à 255.

(a)

(b)

(c)

FIG. 3.24 - Images de bandes verticales de niveaux de gris ayant servi à la construction de fonction de la sensibilité de l'oeil au contraste. a) Bandes contrastées avec un  $C = 4$ . b)  $C = 8$ . c)  $C = 20$ .

Pour une valeur  $C$  qui varie de 1 à 20, une série de 20 images est obtenue. L'observation de l'ensemble de cette série montre que trois zones sont généralement distinguées. La première et la troisième zone sont pratiquement homogènes, c'est à dire que les frontières entre bandes d'une zone sont invisibles par un observateur humain. Par contre les bandes de la deuxième zone se distinguent plus facilement.

Les niveaux de gris de la première et de la dernière bande de la deuxième zone de chaque image (notés respectivement  $l_1$  et  $l_2$ ), ainsi que la valeur de  $C$  associée sont prélevés. En fonction de ces trois grandeurs, on obtient deux valeurs de la fonction  $\Delta[l]$  (fonction des contrastes à peine perceptible par l'oeil humain):  $\Delta[l_1] = C$  et  $\Delta[l_2] = C$ . Cette opération est répétée pour l'ensemble de la série. La perception du contraste entre un point et son voisinage est dépendante de son niveau de gris  $l$ , et de la différence  $C$  des niveaux de gris de ses voisins proches. Une analyse plus fine a été effectuée pour déterminer une fonction de sensibilité au contraste intégrant trois niveaux: contraste fort, moyen, faible, comme le montre la figure 3.25. L'introduction de cette fonction permet de tenir compte de la visibilité des points par rapport à leurs voisins. Elle favorise le regroupement des points de luminances visuellement proches.

FIG. 3.25 - *Fonction de perception visuelle*

En fonction du nombre de critères vérifiés, les fréquences d'apparition  $h_c(i)$  des niveaux de gris  $I(x, y)$  sont modifiées de la manière suivante:

$$\begin{array}{ll}
 h_c(i) = h_c(i) + p_4 & \text{si } (C_1 \text{ et } C_2 \text{ et } C_3 \text{ et } C_4) \text{ sont vrais;} \\
 h_c(i) = h_c(i) + p_3 & \text{si } (C_1 \text{ et } C_2 \text{ et } C_3) \text{ ou } (C_1 \text{ et } C_2 \text{ et } C_4) \\
 & \text{ou } (C_1 \text{ et } C_3 \text{ et } C_4) \text{ ou } (C_2 \text{ et } C_3 \text{ et } C_4) \text{ est vrai;} \\
 h_c(i) = h_c(i) + p_2 & \text{si } (C_1 \text{ et } C_2) \text{ ou } (C_1 \text{ et } C_3) \text{ ou } (C_1 \text{ et } C_4) \\
 & \text{ou } (C_2 \text{ et } C_3) \text{ ou } (C_2 \text{ et } C_4) \text{ ou } (C_3 \text{ et } C_4) \text{ est vrai;} \\
 h_c(i) = h_c(i) + p_1 & \text{si } C_1 \text{ ou } C_2 \text{ ou } C_3 \text{ ou } C_4 \text{ est vrai;} \\
 h_c(i) = h_c(i) & \text{sinon.}
 \end{array}$$

$p_1, p_2, p_3$  et  $p_4$  sont des constantes telles que  $p_1 < p_2 < p_3 < p_4$ . En tenant compte à la fois des relations spatiales entre points et de leur visibilité, cette transformation permet de

mettre en évidence les principaux modes de l'histogramme.

Avant de déterminer les principaux pics de l'histogramme, une opération de lissage lui est appliquée. Celle-ci est nécessaire pour éliminer les irrégularités de l'histogramme. On utilise à cet effet un noyau gaussien de taille cinq. Les niveaux de gris de l'histogramme lissé dont les fréquences d'apparition sont inférieures à la moyenne des  $h(i)$  sont ensuite éliminés.

#### • Recherche des pics significatifs

Les pics significatifs des histogrammes résultant des traitements précédents doivent vérifier les trois critères suivants:

- Un niveau de gris  $i$  est un pic de l'histogramme s'il présente une fréquence d'occurrence locale maximale. Ceci revient à vérifier le critère suivant:

$$h_c(i) > h_c(i - 1) \text{ et } h_c(i) > h_c(i + 1) \quad (3.20)$$

- La distance séparant deux pics significatifs doit être supérieure à un seuil  $S_d$ ; sinon on retient comme pic significatif celui dont la fréquence d'occurrence est la plus grande.
- La hauteur d'un pic doit être supérieure à une fraction  $S_h$  du maximum de l'histogramme. Ceci permet d'éliminer les pics de taille relativement faible.

Les valeurs de seuils  $S_d$  et  $S_h$  sont calculées de manière adaptative et représentent respectivement la moyenne des distances des couples de pics successifs et la moyenne des fréquences d'apparition des pics.

La liste des pics valides servira pour la construction de l'histogramme global des maxima locaux. Celui-ci, noté  $H_{glob}$ , présente une forme multi-modale bien prononcée. Notons également que l'allure globale de cet histogramme peut différer de celle de l'histogramme de l'image. Ceci est dû au fait que dans les zones sombres et claires de l'image, le regroupement des pixels est effectué avec une tolérance importante (le seuil  $\Delta[I(i, j)]$  est élevé), ce qui permet plus probablement de vérifier simultanément les critères  $C_1, C_2, C_3$  et  $C_4$  et ainsi d'augmenter les fréquences des niveaux de gris avec les pas les plus élevés ( $p_3$  et  $p_4$ ).

Les figures 3.27 et 3.29 illustrent respectivement l'application de cette procédure aux images synthétique (Image Savoie bruitée de la banque d'images du GDR-PRC ISIS) 3.26 et réelle (Image SPOT de la forêt de Paimpont, Bretagne) 3.28. Les résultats expérimentaux obtenus ici correspondent à des valeurs de  $p_1 = 1, p_2 = 2, p_3 = 3, p_4 = 4$ .

(a) (b)

FIG. 3.26 - a) Image Savoie bruitée. b) Son histogramme.

FIG. 3.27 - Histogramme global des maxima locaux  $H_{glob}$  de l'image Savoie bruitée.

(a) (b)

FIG. 3.28 - a) Image SPOT de la forêt de Paimpont. b) Son histogramme.

FIG. 3.29 - Histogramme global des maxima locaux  $H_{glob}$  de l'image de la forêt de Paimpont.

#### 3.3.1.1.2 Classification des niveaux de gris

La classification est obtenue après une procédure de regroupement itérative des niveaux de gris opérant sur l'histogramme  $H_{glob}$ . Le but est de détecter les seuils  $s_k$  les plus significatifs en déplaçant itérativement les valeurs de niveaux de gris vers leurs centres de masse jusqu'à idempotence de l'histogramme. Cette procédure est décrite ci-après.

Soit  $H^{(0)}(i) = H_{glob}(i)$ ,  $i \in [i_{min}, i_{max}]$ , l'histogramme initial à traiter, et soit  $H^{(k)}$ , l'histogramme obtenu à l'itération  $k$ . Les fréquences d'occurrence des niveaux de gris à l'itération  $k + 1$  sont déterminées selon les relations suivantes:

$$H^{(k+1)}(i) = \sum_{s \in V_i} H^{(k)}(s) \delta(i - \mu_k(s)) \quad (3.21)$$

où  $\delta(\cdot)$  représente le symbole de Kronecker. et  $\mu_k(s)$  désigne la moyenne glissante à l'itération  $k$ .

$$\mu_k(s) = \frac{\sum_{s \in V_i} s H^{(k)}(s)}{\sum_{s \in V_i} H^{(k)}(s)} \quad (3.22)$$

$V_i$  représente l'ensemble des  $2m + 1$  niveaux de gris voisins de  $i$ , il est défini par:

$$V_i = \{\max((i - m), i_{min}) \dots \min((i + m), i_{max})\}$$

La relation (3.21) traduit le déplacement d'une partie des masses de  $H^{(k)}(i)$  vers leurs centres de gravité locaux. Dans le cas supervisé, le paramètre  $m$  doit être choisi assez petit (de l'ordre de 5 à 10) si on veut obtenir un nombre important de classes.

Pour rendre la procédure de regroupement automatique (non supervisée), on fait varier  $m$  graduellement jusqu'à ce que des tailles successives fournissent des résultats similaires. L'histogramme final obtenu est constitué de masses ponctuelles. A chaque masse correspond une classe et la position de chaque masse indique le niveau de gris qui sera affecté en tant qu'étiquette à chaque classe. Les seuils de segmentation  $s_k$  sont alors donnés par les moyennes des niveaux de gris déterminées par la position de ces masses.

Les histogrammes de la figure 3.30 illustrent respectivement l'application de cette classification aux histogrammes globaux de l'image synthétique (figure 3.26) et réelle (figure 3.28). Les pics importants représentent les masses finales résultant du regroupement. Dans tous les résultats expérimentaux présentés dans cette section, la valeur du paramètre  $m$  varie graduellement de 3 à 9 par pas de 3.

(a)

(b)

FIG. 3.30 - Masses ponctuelles après groupement des masses de  $H_{glob}$ . (a) de l'image Savoie bruitée. (b) l'image de la forêt de Paimpont.

### 3.3.1.1.3 Résultats expérimentaux

Nous avons expérimenté cette méthode sur un échantillon d'images tests couvrant diverses applications. Les figures 3.31 à 3.33 présentent les résultats de multi-seuillage obtenus sur des images provenant de différentes banques d'images. L'ensemble de ces résultats a été obtenu avec le même jeu de paramètres. Pour chacune des images, en plus du résultat du multi-seuillage, sont présentés respectivement de gauche à droite, son histogramme, l'histogramme global et l'histogramme global après la procédure de regroupement des masses ponctuelles.

FIG. 3.31 - *Image Paimpont. Seuils de segmentation: 44 77 100 123 148 176 195 215*

FIG. 3.32 - *Image IRM. Seuils de segmentation: 17 45 72 94 107 123 153 184 207 222*

FIG. 3.33 - *Image Car. Seuils de segmentation: 15 35 55 80 111 135 155 176 194*

Sur les tableaux 3.1, 3.2 et 3.3 sont indiquées les valeurs de seuils obtenues, en appliquant la méthode de multi-seuillage respectivement aux images Aquitaine et Savoie bruitée de la banque d'images du GdR-PRC ISIS et à l'image Paimpont, pour différentes valeurs de  $m$ . On peut observer une similarité entre les valeurs de seuils obtenues sur l'image originale et l'image pré-traitée.

Au vu des différents histogrammes des images tests considérées dans cette expérimentation, deux remarques peuvent ressortir. Tout d'abord, les histogrammes globaux mettent bien en évidence les principaux modes constituant les histogrammes des images originales,

Taille du paramètre $m$	Valeurs des seuils (image originale)	(image pré-traitée)
10	13 31 60 94 113 129 152 172 190 216 239	18 55 87 110 134 155 173 190 221
12	22 63 116 157 185 210 235	19 49 72 87 109 139 177 220
14	23 64 118 159 183 203 229	21 64 108 139 176 211 232
15	24 48 76 119 159 183 221	22 65 106 135 176 207 228
16	60 119 169 185 220	22 65 104 132 175 208 229
17	61 119 156 180 221	60 112 142 176 220
18	61 120 170 219	61 124 176 220
19	62 120 170 219	61 112 141 175 219
20	62 121 170 215	62 107 136 176 217
21	61 119 169 212	62 112 141 176 216
22	62 120 170 212	60 125 175 213
25	62 120 170	61 124 176

TAB. 3.1 - Comportement de la procédure de regroupement pour l'image Aquitaine en fonction de la taille de la fenêtre de la moyenne,  $m$ .

Taille du paramètre $m$	Valeurs des seuils de segmentation
10, 11, 12, 13	113 142
14, 15, 16, 17, 19	113 141
18	112 140

TAB. 3.2 - Comportement de la procédure de regroupement pour l'image Savoie lissée en fonction de la taille de la fenêtre de la moyenne,  $m$ .

Taille du paramètre $m$	Valeurs des seuils (image originale)	(image pré-traitée)
10	43 75 98 122 139 155 178 199 214	45 77 107 146 176 194 214
12	46 85 113 145 176 207	48 101 137 155 175 205
14	47 85 113 144 175 207	49 84 111 146 173 203
15	47 85 111 144 174 207	49 95 122 147 173 204
16	47 89 116 145 173 205	49 86 113 146 172 202
17	47 89 116 145 194	49 90 117 145 192
18	48 92 124 145 193	49 91 117 145 192
19	48 91 124 183	49 100 119 147 191
20	48 111 162 202	50 92 126 165 202
21	48 111 162 202	50 90 137 180 202
22	49 112 168 205	50 113 166 200
25	50 113 182	51 113 177

TAB. 3.3 - Comportement de la procédure de regroupement pour l'image Paimpont en fonction de la taille de la fenêtre de la moyenne,  $m$ .

notamment lorsqu'ils présentent des chevauchements importants. Ensuite, les regroupements des masses ponctuelles des histogrammes globaux convergent correctement vers les modes représentatifs des classes présentes dans l'image.

Il est généralement difficile de juger de la qualité de la segmentation sans vérité terrain. Cependant on peut facilement constater que les segmentations obtenues par la méthode présentée sont visuellement très corrélées avec les images originales. D'une façon générale, les transitions perceptibles par l'oeil sont pratiquement toutes mises en évidence dans les segmentations, et les éléments visuellement importants dans les images originales sont préservés.

Une étude comparative avec quelques méthodes classiques de multi-seuillage est présentée dans le paragraphe suivant.

#### **3.3.1.1.4 Etudes comparatives et évaluation des méthodes de multi-seuillage**

Nous avons vu dans la section 2.4 que les méthodes d'évaluation peuvent être classées en deux principales catégories.

- La première consiste à évaluer la qualité d'une segmentation sans connaissance de la segmentation idéale. Ici, la segmentation optimale devra être définie à partir des caractéristiques souhaitées des régions.
- La seconde catégorie réalise directement une évaluation en fonction de la segmentation de référence (connue a priori). L'évaluation est obtenue par la mesure de la différence entre deux cartes de régions (ou segmentation).

D'après les comparaisons des différentes méthodes d'évaluation présentées dans [370], celles appartenant au groupe des méthodes empiriques basées sur l'erreur entre résultats et image de référence, surpassent celles s'appuyant sur les caractéristiques des objets segmentés. Dans cette dernière catégorie de méthodes, celles utilisant le critère d'uniformité intra-région est réputée être la meilleure.

Afin d'évaluer la méthode de multi-seuillage que nous avons développée et dans le but de comparer ses performances avec les autres méthodes, deux approches d'évaluation (issues des deux catégories d'évaluation) s'appuyant sur des principes différents ont été testées. Cette démarche est également motivée par la volonté de vérifier la conformité du classement final issu de chacune des deux approches.

- Dans la première, nous avons utilisé la mesure de dissimilarité dont le développement est donné au paragraphe 3.2.3.2.2 (qui peut être considérée comme une approximation d'une distance de Hausdorff moyennée). L'approche consiste à comparer le résultat de la segmentation mis en forme de contours avec la carte des contours de l'image de référence.

- Dans la seconde, nous nous servons de la mesure d'uniformité intra-région similaire à celle développée dans [215] et [319] (voir le paragraphe 2.4.2) qui permettait de comparer les méthodes globales de binarisation.

Notons toutefois que (i) la démarche la plus fiable est celle qui consiste à comparer le résultat de la segmentation à une segmentation de référence, produite manuellement par un opérateur humain, connaissant parfaitement le contexte; (ii) il est préférable d'effectuer l'évaluation sur des images naturelles afin de mieux juger la qualité des résultats.

Pour comparer les résultats de segmentation, nous avons choisi deux images tests pour lesquelles nous disposons de la vérité de terrain. La première est une image synthétique de la banque d'images GdR-PRC ISIS. La seconde est une image satellitaire qui représente la forêt de Paimpont.

Notons que les méthodes que nous avons passées en revue dans la section 2.2.1 ne sont pas toutes automatiques. Certaines, nécessitent l'introduction de connaissances a priori sur la scène à segmenter comme le nombre de classes qu'on souhaite obtenir, et d'autres, la définition de paramètres de contrôle tels que le nombre de voisins et la taille de la fenêtre d'analyse pour l'obtention des seuils optimaux. Les méthodes évaluées sont les suivantes:

La méthode basée sur l'analyse discriminante [299] AD;

La méthode basée sur les matrices de co-occurrence de Chanda MC1 [57] et de Weszka MC2;

La méthode des matrices de transition [117] MT;

La méthode basée sur la théorie d'information de Kapur [200] Entropie;

La méthode des matrices de longueurs de plages [130] LdP;

La méthode préservant les moments [333] MOM;

La méthode basée sur sur la transformation d'histogramme [195] THi;

La méthode des nuées dynamiques [111] [113] K-means;

La méthode des centres mobiles [354] [323] CM;

Dans la suite, nous présenterons quelques tests d'évaluation effectués sur un jeu de résultats de segmentation issus de la méthode de multi-seuillage développée (méthode MS) et des méthodes de multi-seuillage global sus-citées. Quelques unes de ces méthodes se trouvent dans la banque de programmes SIMPA du GdR-PRC ISIS, d'autres ont été développées sur des plateformes telles que Khoros et Xite.

## Evaluation avec connaissance de la vérité terrain

### • Image synthétique

Ce paragraphe présente les résultats de comparaison effectuée sur l'image Savoie bruitée aditivement par un bruit blanc gaussien (Fig. 3.34(a)) (image synthétique de la banque d'images du GdR-PRC ISIS).

Les figures 3.35 et 3.36 illustrent les résultats obtenus par les différentes techniques sus-citées, appliquées à l'image Savoie après pré-traitement.

Afin de rendre la comparaison cohérente, les paramètres d'entrée des différentes méthodes de segmentation ont été choisis de manière à avoir un nombre de classes proche de la vérité de terrain et la meilleure similarité suivant la dissimilarité utilisée avec cette vérité de terrain. Toutefois, pour certaines méthodes, il n'a pas été possible d'obtenir le nombre de classes souhaité.

Le tableau 3.4 présente les seuils obtenus pour chacune de ces méthodes.

Dans le tableau 3.5 sont présentées les différences entre les cartes de contours issues de chaque méthode et la carte de contours de l'image de référence (Fig. 3.34(b)).

En examinant les résultats obtenus par les différentes techniques de multi-seuillage sur l'image Savoie, on peut remarquer que les méthodes THi, MT et LdP donnent de moins bons résultats par rapport aux autres méthodes testées.

Cette observation est confirmée par les mesures de dissimilarité données dans le tableau 3.5.

Il est difficile de juger de la qualité des résultats visuellement entre les méthodes restantes. Par contre, si on se réfère aux mesures objectives données par la mesure de dissimilarité employée, les méthodes MC2, MC1, AD se placent parmi les premières.

La méthode que nous avons développée se situe en quatrième position mais rappelons qu'elle est non supervisée, ce qui n'est pas le cas pour les trois premières méthodes qui ont nécessité un réglage de leur paramètres pour obtenir les meilleurs résultats.

(a) (b)

FIG. 3.34 - a) Image Savoie bruitée de la banque d'image du GDR TdSI. b) Image Référence Savoie Rf.

(a) (b) (c)

(d) (e) (f)

FIG. 3.35 - Résultats de segmentation de l'image Savoie bruitée, obtenus par les différentes méthodes de multi seuillage. a) Résultat de segmentation obtenu par la méthode développée. b) Méthode THi. c) Méthode AD. d) Méthode des Kmeans. e) Méthode MC1 avec  $NV = 20$ . f) Méthode MOM avec  $NC = 3$ .

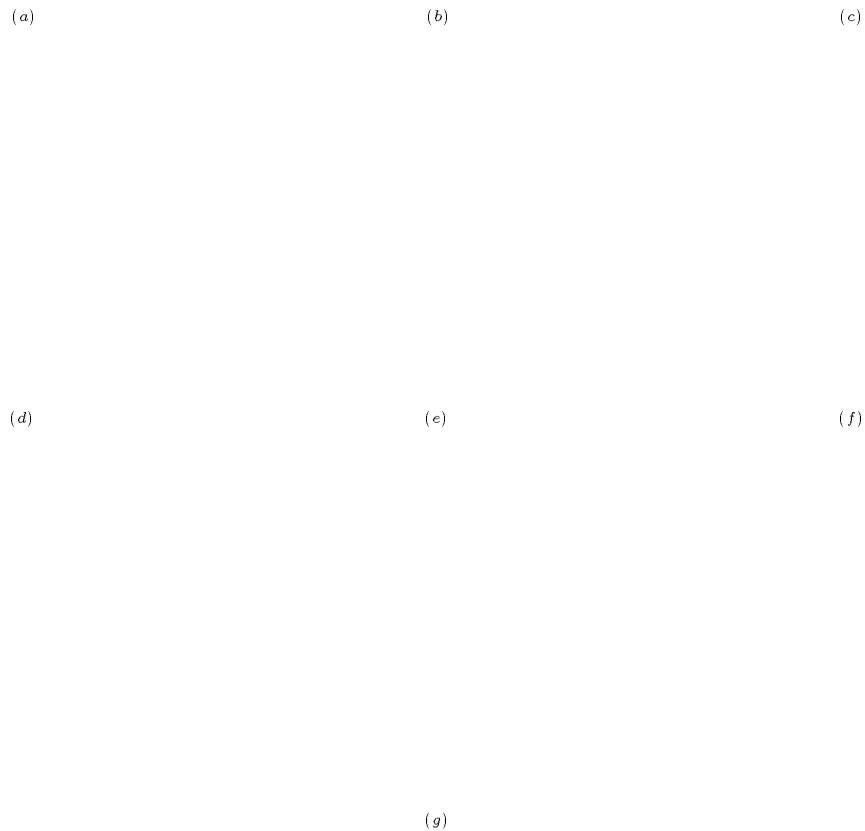


FIG. 3.36 - Résultats de segmentation de l'image Savoie bruitée, obtenus par les différentes méthodes de multi seuillage. a) Méthode MT. b) Méthode d'entropie avec  $NV = 2$ . c) Méthode d'entropie avec  $NV = 6$ . d) Méthode MC2 avec  $NV = 1$ . e) Méthode MC2 avec  $NV = 8$ . f) Méthode LdP  $NV = 23$ . g) Méthode CM.

	Paramètres	Nombre de classes	Seuils détectés
Méthode MS développée		3	113 142
Méthode AD	$NC = 3$	3	113 143
Méthode Kmeans	$NC = 3$	3	113 142
Méthode MC1	$NV = 10$	4	111 143 168
	$NV = 20$	3	111 143
Méthode MT	$NV = 8$	4	107 134 150
	$NV = 10$	3	107 143
Méthode d'entropie	$NV = 2$	4	112 115 144
	$NV = 3$	3	115 144
Méthode MC2	$NV = 1$	3	111 143
Méthode LdP	$NV = 23$	3	127 159
Méthode MOM	$NC = 3$	3	108 134
Métho THi	$NS = 2$	3	124 127
Méthode CM	$W = 15$	3	112 140

TAB. 3.4 - Application des différentes méthodes de multi-seuillage sur l'image Savoie.  $W$ : taille de fenêtre,  $NC$ : nombre de classes,  $NV$ : nombre de voisins,  $NS$ : nombre de seuils.

	Paramètre	$D_M(Rf, RS_i)$	$D_M(RS_i, Rf)$	Dissimilarité
Méthode MC2	$NV = 1$	0.002792	0.008679	0.005735
Méthode MC1	$NV = 10$	0.002813	0.010639	0.006726
	$NV = 20$	0.002815	0.008969	0.005892
Méthode AD	$NC = 3$	0.002814	0.010686	0.006750
Méthode MS développée		0.002814	0.011080	0.006947
Méthode Kmeans	$NC = 3$	0.002814	0.011080	0.006947
Méthode MOM	$NC = 3$	0.002820	0.011526	0.007173
Méthode CM	$W = 15$	0.002814	0.012869	0.007842
Méthode d'entropie	$NV = 2$	0.002809	0.014170	0.008489
Méthode de MT	$NV = 8$	0.003004	0.035891	0.019448
	$NV = 10$	0.003330	0.038455	0.020892
Méthode LdP	$NV = 23$	0.004126	0.041829	0.022977
Métho THi	$NS = 2$	0.004568	0.048295	0.026432

TAB. 3.5 - Dissimilarités calculées entre les différents résultats de multi-seuillage  $RS_i$  de l'image Savoie et la segmentation de référence  $Rf$ .  $W$ : taille de fenêtre,  $NC$ : nombre de classes,  $NV$ : nombre de voisins,  $NS$ : nombre de seuils.

- Image réelle

Dans ce paragraphe nous présentons les résultats de comparaisons effectuées sur l'image Paimpont (figure 3.37). Pour cette image nous disposons de deux vérités de terrain. La première notée  $V_1$  représente une classification de l'image originale en 20 classes. Dans la seconde (carte  $VT$ ) certaines classes de  $V_1$  ont été regroupées pour obtenir un nombre plus réduit de classes (12).

Nous présentons dans les figures 3.38 et 3.39 les résultats obtenus par les différentes techniques de multi-seuillage.

Le tableau 3.6 présente les seuils obtenus pour chacune de ces techniques, ainsi que les valeurs des paramètres utilisés pour avoir un nombre de classes équivalent à celui souhaité.

Dans le tableau 3.7 sont présentées les dissimilarités normalisées entre les cartes de contours issues de chacune des méthodes de multi-seuillage et la carte de contours de référence  $V_1$

Sur la deuxième colonne sont rappelées les valeurs de paramètres utilisées pour obtenir la segmentation la plus similaire à la vérité de terrain.

Nous présentons à titre indicatif sur la troisième colonne (resp. quatrième colonne) la dissimilarité dans la direction vérité de terrain vers résultat de segmentation (resp. résultat de segmentation vers vérité de terrain).

Vu les résultats de multi-seuillage présentés dans les figures 3.38 et 3.39, il est difficile de classer les performances des différentes méthodes.

Par ailleurs, on peut observer en s'aidant du tableau 3.7 que les méthodes qui ont donné un résultat convenable dans le cas de l'image Savoie bruitée, ne fournissent pas les meilleurs résultats pour l'image Paimpont. Cette remarque s'applique particulièrement aux résultats obtenus par la méthode MC1 et celle MC2. Ceci peut en partie s'expliquer par le fait que ces méthodes supposent une uniformité de la distribution des niveaux de gris ce qui n'est pas le cas de l'image Paimpont.

(a) (b)

FIG. 3.37 - (a) Image de la forêt de Paimpont masquée. (b) Contours vérité de terrain de la forêt de Paimpont. Image VT.

(a) (b) (c)

(d) (e) (f)

FIG. 3.38 - Résultats de segmentation de l'image Paimpont obtenus par les différentes méthodes de multi seuillage. a) Résultat de segmentation obtenus par la méthode développée. b) Méthode THi. c) Méthode AD. d) Méthode des Kmeans. e) Méthode MC1 avec  $NV = 1$ . f) Méthode MC1 avec  $NV = 2$ .

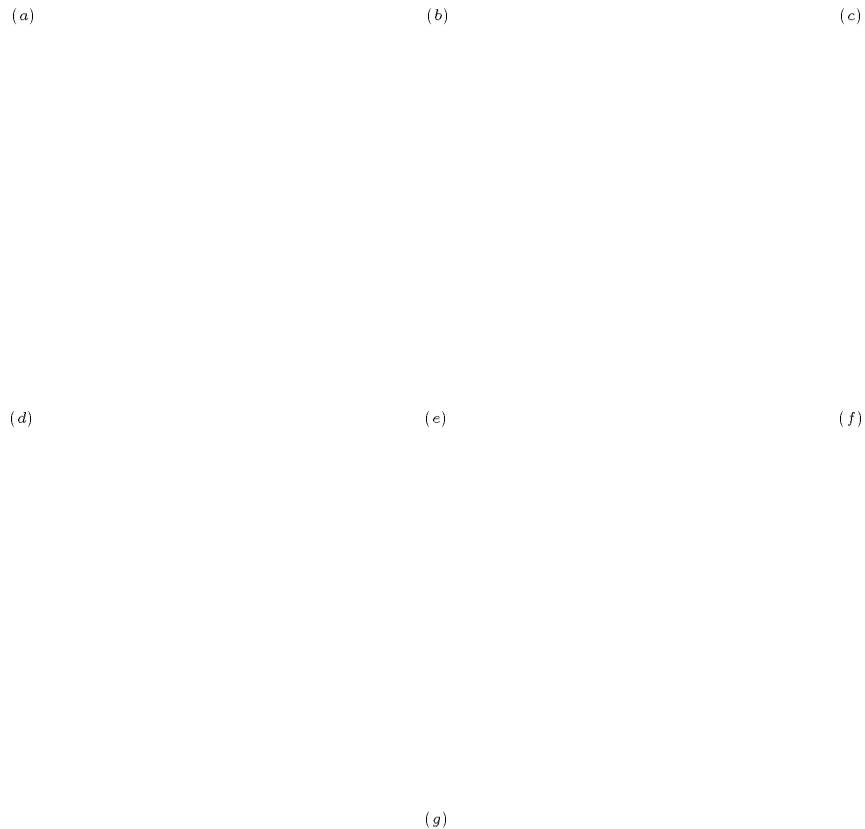


FIG. 3.39 - Résultats de segmentation de l'image Paimpont obtenus par les différentes méthodes de multi seuillage. a) Méthode MT. b) Méthode d'entropie avec  $NV = 5$ . c) Méthode d'entropie avec  $NV = 6$ . d) Méthode MC2 avec  $NV = 1$ . e) Méthode MC2 avec  $NV = 2$ . f) Méthode LdP. g) Méthode CM.

	Paramètres	Nombre de classes	Seuils détectés
Méthode MS développée		7	32 83 115 138 157 192
Méthode AD	$NC = 7$	7	34 77 98 120 140 158
Méthode Kmeans	$NC = 7$	7	10 41 69 87 111 138
Méthode MC1	$NV = 1$	12	50 92 108 112 118 121 147 151 154 164 175
	$NV = 2$	4	50 112 118
Méthode MT	$NV = 8$	7	60 98 108 117 132 161
Méthode d'entropie	$NV = 5$	8	73 90 97 103 124 140 175
	$NV = 6$	6	73 90 97 124 175
Méthode MC2	$NV = 1$	9	86 92 96 100 102 104 129 132
	$NV = 2$	6	86 92 96 102 129
Méthode LdP	$NV = 8$	7	68 81 98 153 217 239
Méthode MOM	$NC = 7$	7	35 66 91 114 138 170
Métho THi	$NS = 6$	7	0 83 103 117 133 139
Méthode CM	$W = 13$	7	11 49 89 123 153 188

TAB. 3.6 - Application des différentes méthodes de multi-seuillage sur l'image Paimpont.  $W$ : taille de fenêtre,  $NC$ : nombre de classes,  $NV$ : nombre de voisins,  $NS$ : nombre de seuils.

	$NC$	$D_M(VT, RS_i)$	$D_M(RS_i, VT)$	Dissimilarité
Méthode AD	7	0.004983	0.004720	0.004852
Méthode MS développée	7	0.005282	0.004746	0.005014
Méthode Kmeans	7	0.005707	0.004618	0.005162
Méthode d'entropie	8	0.005716	0.004670	0.005193
	6	0.007769	0.004633	0.006101
Méthode MT	7	0.005985	0.004713	0.005349
Méthode MOM	7	0.005982	0.004946	0.005464
Méthode LdP	7	0.006111	0.004870	0.005491
Méthode CM	7	0.006335	0.004698	0.005517
Métho THi	7	0.007169	0.004133	0.005652
Méthode MC1	12	0.006382	0.004983	0.005682
	4	0.011785	0.004279	0.008327
Méthode MC2	9	0.007722	0.004590	0.006156
	6	0.007960	0.004561	0.006261

TAB. 3.7 - Dissimilarités calculées entre les différents résultats de multi-seuillage  $RS_i$  de l'image Paimpont et la carte vérité de terrain  $V_1$ .  $NC$ : nombre de classes.

### Evaluation de la segmentation basée sur l'uniformité des niveaux de gris intra-région

Comme il a été déjà signalé, dans cette approche, l'idée de base repose sur le fait qu'une segmentation adéquate devrait produire une image ayant une uniformité intra-région d'autant plus grande que les attributs des éléments (niveau de gris) de la région sont similaires.

L'uniformité d'une caractéristique sur une région peut être calculée sur la base de la variance des caractéristiques évaluées à chaque pixel appartenant à cette région.

En particulier, pour une image à niveau de gris  $I(x, y)$ , soient  $R_i$  la  $i^{ieme}$  région de l'image segmentée et  $A_i$  l'aire de  $R_i$ , la mesure de l'uniformité du niveau de gris ( $GU_{reg}$ ) de  $I(x, y)$  s'écrit:

$$GU_{reg} = \sum_i \sum_{(x,y) \in R_i} [I(x, y) - \frac{1}{A_i} \sum_{(x,y) \in R_i} I(x, y)]^2 \quad (3.23)$$

Partant de cette définition, une mesure d'uniformité normalisée ( $GU_{regnorm}$ ) a été proposée par Sahoo et al. [319]:

$$GU_{regnorm} = 1 - \frac{GU_{reg}}{GU_{max}} \quad (3.24)$$

où  $GU_{max}$  est un facteur de normalisation<sup>2</sup>.

$$GU_{max} = \text{Nombre de lignes} * \text{Nombre de colonnes} * (128)^2 \quad (3.25)$$

Le traitement algorithmique de cette méthode d'évaluation peut être assez long dans le cas d'images complexes pouvant comporter un nombre important de régions. Ainsi, dans le même esprit que la méthode d'évaluation précédente, le calcul est opéré sur chaque niveau de gris et non sur chaque région. La mesure de l'uniformité par niveau de gris ( $GU_{gris}$ ) s'écrit donc:

$$GU_{gris} = \sum_i \sum_{(x,y) \in g_i} [I(x, y) - \frac{1}{A_i} \sum_{(x,y) \in g_i} I(x, y)]^2 \quad (3.26)$$

où  $g_i$  est un niveau de gris de l'image segmentée et  $A_i$  est le nombre de points de niveaux de gris  $g_i$ .

Comme précédemment, la valeur normalisée de  $GU_{gris}$  est:

$$GU_{grisnorm} = 1 - \frac{GU_{gris}}{GU_{max}} \quad (3.27)$$

Nous présentons dans le tableau 3.8 (resp. 3.9) la mesure d'uniformité par région (resp. par niveau de gris) sur les résultats de segmentation de l'image Savoie bruitée.

---

2. La plus mauvaise mesure de l'uniformité des niveaux de gris intra-région est obtenue quand l'image est composée de deux régions identiques dont l'une a le niveau de gris 0 et l'autre 255 alors que le résultat final aurait dû être une image uniforme avec une seule région ayant un niveau de gris égal à 128.

	Paramètres	$GU_{reg}$ normalisée
Méthode AD	7	0.996703
Méthode MS développée		0.996494
Méthode MOM	7	0.996444
Méthode d'entropie	4	0.996427
Méthode Kmeans	7	0.996424
Méthode CM	13	0.996325
Méthode MT	5	0.996256
Méthode MC1	4	0.995918
Méthode MC2	2	0.995405
Méthode LdP	6	0.994105
Métho THi	7	0.992286

TAB. 3.8 - Mesures selon l'uniformité des niveaux de gris par région. Image Savoise bruitée.

	Paramètres	Nbre de classe	$GU_{gris}$ normalisée
Méthode AD	3	3	0.995698
Méthode MS développée		3	0.995619
Méthode MOM	3	3	0.995613
Méthode Kmeans	3	3	0.995520
Méthode MT	5	3	0.995233
Méthode Entropie	4	3	0.995103
Méthode MC1	4	3	0.995190
Méthode CM	15	7	0.995008
Méthode MC2	2	3	0.994449
Méthode LdP	6	4	0.993561
Métho THi	2	3	0.990861

TAB. 3.9 - Mesures selon l'uniformité des niveaux de gris par niveau de gris. Image Savoise bruitée.

	Paramètres	$GU_{reg}$ normalisée
Méthode AD	7	0.998582
Méthode MS développée		0.998004
Méthode MC1	1	0.997772
Méthode Kmeans	7	0.997537
Méthode LdP	7	0.997352
Métho THi	7	0.997314
Méthode MT	8	0.996983
Méthode CM	13	0.996648
Méthode MOM	7	0.996581
Méthode MC2	1	0.964435
Méthode d'entropie	5	0.986240

TAB. 3.10 - Mesures selon l'uniformité des niveaux de gris par région. Image Paimpont.

	Paramètres	Nbre de classe	$GU_{gris}$ normalisée
Méthode AD	7	7	0.998374
Méthode MS développée		7	0.997838
Méthode MOM	7	7	0.997712
Méthode MC1	1	8	0.997433
Méthode Kmeans	7	7	0.997270
Méthode LdP	7	9	0.997051
Métho THi	6	7	0.996897
Méthode CM	13	7	0.995648
Méthode MT	8	9	0.994289
Méthode d'entropie	5	10	0.979670
Méthode MC2	1	7	0.959823

TAB. 3.11 - Mesures selon l'uniformité des niveaux de gris par niveau de gris. Image Paimpont.

Comme précédemment, nous avons procédé à la mesure d'uniformité par région et par niveau de gris sur les résultats de segmentation de l'image Paimpont. Les tableaux 3.10 et 3.11 présentent respectivement ces deux mesures.

### Conclusion sur l'évaluation

Pour comparer les différentes méthodes de multi-seuillage, deux images test et deux approches d'évaluation s'appuyant sur des principes différents ont été examinées. Dans la première approche d'évaluation, nous avons utilisé la mesure de dissimilarité (cf. 3.2.3.2.2). Cette approche consiste à comparer le résultat de la segmentation obtenu sous forme de contours avec la carte des contours de l'image de référence. La seconde approche fait appel à la notion d'uniformité des niveaux de gris, et est adaptée à l'évaluation de la segmentation en l'absence de vérité de terrain.

L'examen des différentes mesures selon les deux approches d'évaluation appliquées aux images tests, montre une certaine cohérence du classement final des méthodes comparées.

Les résultats de comparaison basés sur les critères d'uniformité ont montré que la mé-

thode basée sur l'analyse discriminante (AD) avec celle que nous avons développée donnent les meilleurs résultats, pour l'image satellitaire. La méthode basée sur la préservation des moments (MOM) ainsi que la méthode des Kmeans donnent également de bons résultats.

### 3.3.1.1.5 Conclusion

Nous avons développé dans cette section un algorithme de segmentation basé sur le multi-seuillage d'un histogramme issu d'une analyse locale de l'image.

L'originalité de la méthode de multi-seuillage réside, d'une part, dans la prise en compte des propriétés liées à la perception visuelle, et, d'autre part, dans l'introduction de manière indirecte des statistiques d'ordre deux, sous une forme unidimensionnelle, afin d'obtenir une division efficace. Notons enfin, que la technique de multi-seuillage tient compte des propriétés locales d'homogénéité de chaque histogramme et ne demande aucune connaissance a priori sur l'image à segmenter.

Les résultats de cette procédure se révèlent corrects dans le cas des images composées d'un certain nombre de régions bien contrastées, avec une faible variation des luminances à l'intérieur d'une même région. Mais les résultats s'avèrent peu intéressants dès que les éléments de l'image sont variés et nombreux, dans le cas des images acquises avec des conditions d'éclairage non uniforme ainsi qu'en présence de zones fortement texturées.

Dans ce dernier cas, l'intérêt de cette première analyse réside dans l'extraction de l'information susceptible de ne contenir, compte tenu de l'aspect textuel de l'image, que les éléments dont l'influence est la plus significative pour la caractérisation des différentes régions de l'image. Ceci permet une première simplification des données avant de procéder au traitement complémentaire présenté dans la section 3.3.1.2.

### 3.3.1.2 Extraction des textures par caractérisation et fusion des régions multi-seuillées

#### 3.3.1.2.1 Introduction

En général, les résultats peu satisfaisants dans le cas de l'approche par extraction de textures, sont en grande partie dus d'une part, aux problèmes d'extraction d'informations caractéristiques permettant d'augmenter la séparation entre les différents modèles de texture et d'autre part, à la fluctuation des niveaux de gris de zones homogènes.

Les résultats de division obtenus par la méthode de multi-seuillage développée ci-dessus sont satisfaisants dans la mesure où les régions homogènes au sens des niveaux de gris ont été correctement extraites. Le problème demeure dans le cas des régions texturées et ayant des luminances éloignées. Une analyse plus globale est alors nécessaire.

En considérant qu'une texture est une structure disposant de certaines propriétés spatiales homogènes et invariantes par translation, nous avons utilisé pour la caractériser les étiquettes issues de la procédure de multi-seuillage et leurs fréquences d'occurrence. Dans la suite du document nous utiliserons le terme "image des étiquettes" pour désigner le résultat de la procédure de multi-seuillage.

Le problème de segmentation est ainsi ramené à la localisation des différentes régions de l'image des étiquettes ayant un comportement semblable au sens des étiquettes et de leurs fréquences d'apparition.

La segmentation consiste en:

- l'extraction de caractéristiques locales dans les images; ces caractéristiques forment un vecteur de paramètres.
- le regroupement des vecteurs dans l'espace de paramètres est ensuite effectué en utilisant un algorithme de classification (de type "plus proches voisins").

Nous présentons, ci-après, les algorithmes de fusion développés ainsi que les attributs utilisés pour décrire la texture.

### **3.3.1.2.2 Localisation grossière des textures**

Pour localiser les textures dans l'image nous avons expérimenté plusieurs méthodes basées sur l'utilisation des histogrammes généralisés des niveaux de gris issus de la procédure du multi-seuillage. Nous en décrivons, ci-dessous, deux méthodes. Dans la première, nous exploitons les histogrammes du premier ordre [187]. La seconde méthode introduit les histogrammes d'ordre deux. Nous présentons ensuite quelques résultats obtenus sur des images synthétiques et réelles comprenant des régions uniformes et texturées.

#### **Méthode à base de statistiques d'ordre un**

Le principe général pour localiser les zones texturées de l'image s'appuie sur un processus de fusion de régions. L'algorithme comprend trois phases:

- **Phase 1]** l'image est découpée en blocs rectangulaires qui se chevauchent;
- **Phase 2]** les statistiques locales sont, ensuite, calculées à l'intérieur de chacun des blocs de l'image des étiquettes et de l'image originale correspondante;
- **Phase 3]** une fusion des blocs minimisant un critère de similarité est effectuée.

Nous détaillons chacune de ces étapes dans les paragraphes suivants.

**Phase 1] découpage de l'image:** Le découpage initial de l'image est réalisé indépendamment des données, et consiste à partitionner l'image en blocs rectangulaires. La taille des blocs doit être assez grande pour que le calcul des statistiques soit significatif. Cette taille doit malgré tout rester raisonnable afin de calculer les caractéristiques de façon locale.

Le choix de la taille du bloc est inhérent à tous les problèmes d'analyse d'images. Il dépend de la résolution de la texture à analyser. Les tests effectués sur des images aériennes et sur des images de scènes d'extérieur montrent qu'une taille de 8 x 8 pixels amène à un trop grand nombre de blocs sans une amélioration sensible de la qualité des résultats de fusion. Les résultats obtenus avec des blocs de 32 x 32 pixels ne sont pas suffisamment précis. Dans le cas des images satellitaires, un bon compromis peut être obtenu avec des blocs de taille de 16 x 16 pixels qui se chevauchent.

**Phase 2] attributs de textures:** Pour la caractérisation de la texture d'un bloc, deux possibilités ont été étudiées. Dans la première, elle est caractérisée par les étiquettes et leurs fréquences d'occurrence. Dans la seconde, la moyenne des niveaux de gris  $\mu_e$  dans l'image des étiquettes et les deux premiers moments  $\mu_i$  et  $\sigma_i$  (moyenne et écart-type) des niveaux de gris calculés sur l'image originale, sont utilisés.

$$\mu_e = \sum_{l \in W_e} lh_e(l) \quad (3.28)$$

où

$W_e$  correspond à la fenêtre d'analyse associée à l'image des étiquettes,

$h_e$  correspond à l'histogramme des étiquettes de  $W_e$ .

et

$$\mu_i = \sum_{g \in W_i} gh(g) \quad (3.29)$$

$$\sigma_i = \sqrt{\sum_{g \in W_i} (g - \mu_i)^2 h(g)} \quad (3.30)$$

où

$W_i$  correspond à la fenêtre d'analyse associée à l'image originale,

$h$  correspond à l'histogramme des niveaux de gris de  $W_i$ .

Notons que d'autres attributs peuvent être extraits à partir des histogrammes d'ordre un, parmi lesquels on peut citer l'entropie, l'énergie,...

**Phase 3] fusion de régions minimisant un critère de similarité:** L'algorithme examine l'image originale et celle des étiquettes bloc par bloc selon un schéma séquentiel de

parcours. Pour chaque bloc on cherche dans son voisinage  $V$  selon la figure (3.40) le bloc qui minimise un critère de similarité. Tous les blocs connexes vérifiant ce critère (appartenant à la même texture) sont alors regroupés pour former la région.

FIG. 3.40 - *Voisinage utilisé pour la segmentation en régions. (Avec chevauchement spatial).*

Le voisinage  $V$  utilisé est constitué des  $c$  blocs précédents et des  $2c + 1$  blocs des  $r$  lignes supérieures.  $c$  et  $r$  sont des variables calculées en fonction de la taille du bloc d'observation.

Le problème de la distance de similarité entre deux textures n'est toujours pas résolu en théorie. Généralement se sont des distances classiques issues de la géométrie euclidienne qui sont utilisées. On peut à ce propos noter les travaux de Satini et Jain qui proposent certaines propriétés nécessaires pour les mesures de similarité et les discutent sur des distances classiques [315]. Dans notre cas, pour mesurer le degré de similarité entre le bloc examiné, noté  $B^c$ , et un de ses blocs voisins, noté  $B^v$ , nous avons expérimenté deux critères. Le premier utilise la distance de Kolmogorov et le deuxième fait intervenir la somme des différences absolues des moyennes des niveaux de gris des étiquettes ( $\mu_e^c$  et  $\mu_e^v$ ) et les deux premiers moments ( $\mu_i^x$  et  $\sigma_i^x$ , ( $x = c, v$ )) calculés sur les points de l'image originale.

Pour deux histogrammes d'étiquettes  $h_{e1}$  et  $h_{e2}$ , la distance de Kolmogorov est donnée par

$$D_{(h_{e1}, h_{e2})}^K = \sum_{l=0}^{N_E-1} |h_{e1}(l) - h_{e2}(l)| \quad (3.31)$$

$N_E$  représente le nombre des étiquettes. Une propriété remarquable de cette distance est son invariance par toute transformation strictement monotone des données. Le critère de similarité entre deux blocs voisins  $B^c$  et  $B^v$  est alors donné par

$$C_h = (D_{(h_e^c, h_e^v)}^K < s_h) \quad (3.32)$$

Les deux blocs sont considérés comme similaires si la distance entre leurs histogrammes  $h_e^c$  et  $h_e^v$  est inférieure au seuil  $s_h$ . Ce seuil est réglé empiriquement et choisi de l'ordre de 20% à 40% de la surface du bloc. Dans le cas où aucun bloc adjacent n'est similaire au bloc courant, on entreprend la construction d'une nouvelle région dont celui-ci est le premier élément. Ce processus est itéré jusqu'à balayage complet de l'image des étiquettes.

Le deuxième indice que nous avons défini est donné par l'expression suivante :

$$\delta_{h1}(B^c, B^v) = |\mu_e^c - \mu_e^v| + |\mu_i^c - \mu_i^v| + |\sigma_i^c - \sigma_i^v| \quad (3.33)$$

Comme dans le premier cas, deux blocs sont considérés comme similaires si leur indice de similarité est inférieur à un seuil  $s_h$ .

L'utilisation des histogrammes des valeurs des étiquettes et des deux premiers moments des niveaux de gris s'avère fiable dans certains cas de figures. Toutefois, ces statistiques ne sont pas suffisantes pour caractériser entièrement la texture car elles ne tiennent pas compte de l'arrangement des étiquettes et des relations spatiales qui les gèrent. Cette limitation engendre une faiblesse de leur pouvoir de discrimination notamment dans le cas de textures à caractère régulier. On peut illustrer ceci par l'exemple de la figure 3.41, où sont représentés des échantillons de textures différentes ayant des histogrammes d'ordre un identiques.

FIG. 3.41 - *Histogramme identique pour des textures différentes.*

Pour remédier à cet handicap, un recours à des statistiques tenant compte des dépendances spatiales des points de l'image s'avère nécessaire afin de conduire à une meilleure discrimination.

### Méthode à base de statistiques d'ordre deux

L'introduction des statistiques d'ordre deux est réalisée grâce aux histogrammes du second ordre (matrices de co-occurrences). Notre second algorithme de fusion de régions repose sur ces matrices calculées sur les niveaux de gris issus de la phase du multi-seuillage. Ceci a pour effet de diminuer considérablement la taille des matrices de co-occurrences. Ainsi, un fort gain en place mémoire est obtenu et les temps mis pour l'analyse des matrices sont plus faibles.

Le principe général est identique à l'algorithme de fusion à base des statistiques d'ordre un précédent. La méthode comprend aussi trois phases. La **phase 1** (découpage de l'image) est la même que précédemment. Les deux autres phases de l'algorithme sont détaillées ci-dessous.

**Phase 2] extraction des paramètres:** Pour chaque bloc de la grille initiale, nous déterminons l'histogramme d'ordre deux. Afin de réduire le coût calculatoire, l'histogramme d'ordre deux est calculé seulement pour les vecteurs de déplacement  $\vec{\delta} = (\delta_i, \delta_j) = (0, 1)$  (déplacement horizontal), et  $(1, 0)$  (déplacement vertical).

Pour l'extraction des paramètres, deux possibilités sont également envisageables: l'utilisation des matrices de co-occurrence des étiquettes ou bien l'extraction d'attributs sur celles-ci. Dans ce dernier cas, comme nous l'avons vu au paragraphe 2.2.3.2.2, un grand nombre de mesures peut être défini à partir de la matrice de co-occurrence  $P(l_0, l_1; \vec{\delta})$ , pour caractériser les régions texturées de l'image. En raison du coût de calcul élevé, seuls quelques uns de ces attributs sont généralement utilisés. Il s'agit de: l'uniformité de l'énergie, l'inertie (le contraste), la corrélation, l'homogénéité locale, et l'entropie.

Ces paramètres paraissent souvent difficilement interprétables si l'on se réfère à l'aspect visuel de la texture analysée. On peut cependant indiquer que:

- Une image à texture uniforme a une uniformité égale ou proche de l'unité.
- L'inertie étant fonction des différences  $(l_0 - l_1)$ , elle est d'autant plus faible que les valeurs élevées de  $P(l_0, l_1; \vec{\delta})$  sont regroupées sur la diagonale de la matrice.
- L'entropie d'une texture dénote son aspect aléatoire.

L'ensemble des attributs extraits pour les vecteurs de déplacement  $(0, 1)$  et  $(1, 0)$  caractérisent chacun des blocs issus du découpage, et forment ainsi les vecteurs de paramètres des blocs.

**Phase 3] fusion des régions minimisant le critère de similarité:** Pour mesurer le degré de similarité entre blocs nous avons défini un indice faisant intervenir les composantes qui caractérisent chacun des blocs.

Si  $V_c$  représente le vecteur d'attributs du bloc analysé  $B^c$  et  $V_v$  celui d'un bloc voisin  $B^v$ , la similarité des deux blocs est évaluée par le cosinus de l'angle (coefficient de corrélation) entre ces deux vecteurs:

$$\delta_{h2}(B^c, B^v) = \frac{\langle V_c, V_v \rangle}{\|V_c\| \|V_v\|} \quad (3.34)$$

$\langle u, v \rangle$  désigne l'opération du produit scalaire des vecteurs  $u$  et  $v$ .

Comme dans la méthode précédente, deux blocs sont considérés comme similaires si leur indice de similarité est inférieur à un seuil  $s_{h2}$ . Dans le cas où aucun des blocs adjacents n'est similaire au bloc examiné on entreprend la construction d'une nouvelle région dont celui-ci est le premier élément. Ce processus est itéré jusqu'à balayage complet de l'image issue du multi-seuillage. Le seuil de tolérance pour mesurer la similarité est réglé empiriquement en

fonction de la taille des blocs: plus la taille du bloc est petite, plus faible sera la valeur du seuil.

### 3.3.1.2.3 Affinement des frontières par modélisation auto-régressive bidimensionnelle de la texture

Pour déterminer avec plus de précision les contours séparant les différentes textures, nous avons utilisé une deuxième procédure s'appuyant sur la modélisation auto-régressive des régions issues de la fusion précédente.

L'image des étiquettes résultante de la dernière fusion devient la référence spatiale pour la modélisation auto-régressive bidimensionnelle (AR-2D) de la texture. Ainsi, pour chaque étiquette, les paramètres du modèle AR sont estimés selon le procédé décrit au paragraphe 2.2.3.3.1.

Un support demi-plan non-symétrique (DPNS) a été choisi pour les raisons suivantes: d'abord, l'estimation des paramètres par l'approche du maximum de vraisemblance est linéaire et garantit presque sûrement la stabilité du modèle [61] [66]; ensuite, le modèle causal est intéressant pour la détection de rupture de modèle; et enfin, un critère simple est mis en oeuvre afin de déterminer si deux régions connexes présentant des textures homogènes peuvent être fusionnées ou non.

La procédure de fusion liée au modèle AR est effectuée selon le critère suivant.

Soit  $k$  une étiquette de l'image résultant de la première fusion et soit  $T_k$  la règle de décision telle que:

$$T_k = n_k \ln \hat{\sigma}_k^2 + q \ln n_k \quad (3.35)$$

où  $\hat{\sigma}_k^2$  est la variance de l'erreur de prédiction,  $n_k$  est le nombre de pixels de la région de l'étiquette  $k$ , et  $q$  est le nombre de paramètres du modèle AR. Dans notre cas nous avons fixé  $q$  à 24 (moitié d'une fenêtre 7x7).

Pour deux régions connexes d'étiquettes  $i$  et  $j$ , le critère de fusion utilisé est le suivant:

$$T_{i \cup j} < T_i + T_j \quad (3.36)$$

Après cette procédure, à chaque nouvelle étiquette obtenue correspond un vecteur de paramètres composé des paramètres AR, du niveau de gris moyen et de la variance de l'erreur de prédiction linéaire.

Afin d'affiner les frontières des régions constituées dans l'étape précédente, les paramètres AR sont utilisés pour détecter un changement important du comportement de l'erreur de

prédiction selon le principe suivant.

Soient deux régions  $R_i$  et  $R_j$  respectivement étiquetées  $i$  et  $j$ . Lorsque le filtre inverse associé à l'étiquette  $i$  est propagé sur la région  $R_i$ , on obtient théoriquement une moyenne nulle de l'erreur résiduelle et la variance  $\sigma_i^2$  minimale. Le passage du filtre inverse correspondant à l'étiquette  $i$  sur la région  $R_j$  augmente la variance de l'erreur résiduelle et la distribution peut ne plus être centrée. C'est grâce à l'augmentation de la variance que les contours sont détectés.

Toutefois, il se peut que pour une texture donnée la propagation d'un filtre inverse sur une région sur laquelle le modèle AR correspondant à ce filtre inverse n'est pas approprié puisse conduire à une variance plus faible que sur la région  $i$  donnée. Ceci est dû au fait que la variance du bruit résiduel est aussi la variance du processus d'innovation qui conduit à la représentation de la texture à travers le modèle AR.

Pour pallier cette difficulté, on compare simultanément les résultats des erreurs de prédiction  $e_i$  (resp  $e_j$ ) résultant de la propagation du modèle de l'étiquette  $i$  (resp  $j$ ) à l'intérieur de la région  $(R_i \cup R_j)$  composée des deux régions étiquetées  $i$  et  $j$  à l'écart type  $\sigma$  donné par

$$\sigma = \max\{\hat{\sigma}_i, \hat{\sigma}_j\} \quad (3.37)$$

Plus précisément, un pixel  $(m, n)$  appartient au contour de la région  $(R_i \cup R_j)$  si:

$$e_i^2(m, n) > b\sigma^2 \quad \text{et} \quad e_j^2(m, n) > b\sigma^2 \quad (3.38)$$

où  $b$  est un paramètre ajustable (typiquement  $b = 4$ ). Cette règle de décision permet de détecter un nombre important de contours des régions de l'image originale.

### Résultats expérimentaux et conclusion

Nous présentons dans ce paragraphe quelques résultats expérimentaux obtenus sur une image synthétique composée de régions de faibles et de fortes textures extraite de l'album de Brodatz (D24, D29, D38, D68) (figure 3.42(a)). La figure 3.42(b) donne le résultat du multi-seuillage.

Les tests effectués sur des images de textures naturelles de Brodatz montrent qu'il n'y a pas de différence notable entre les résultats obtenus par les méthodes des statistiques d'ordre un en utilisant les deux indices de similarité.

Des illustrations de la première méthode de segmentation par fusion de régions avec le premier indice de similarité (d'histogrammes d'étiquettes), appliquée aux images des figures 3.42(a) et 3.42(b) sont données figures 3.42(c) et 3.42(d).

Les résultats de la fusion, présentés sont respectivement obtenus avec des tailles des blocs de 32 x 32 et 16 x 16 en effectuant un balayage avec chevauchement spatial de ceux-ci.



FIG. 3.42 - (a) Image originale composée des textures D24, D29, D38 et D68. (b) Résultat du multi-seuillage de l'image de textures. (c) Résultat de la segmentation par fusion de régions après un filtrage médian. avec une fenêtre de taille  $32 \times 32$ , et un seuil de fusion égal à 35% de la surface de la fenêtre. (d) Avec une fenêtre de taille  $16 \times 16$ , et un seuil de fusion égal à 15% de la surface de la fenêtre.

Une illustration de la seconde méthode de segmentation par fusion de régions (statistiques d'ordre deux) appliquée aux images des figures 3.42(a) et 3.42(b) est donnée figure 3.43 (a). Pour cette expérimentation, nous avons choisi d'utiliser des blocs de  $16 \times 16$  pixels, ce qui paraît être un bon compromis entre les arguments précédemment cités. En effet, une taille de  $8 \times 8$  pixels amène à un nombre de blocs important et donc à des temps de calcul élevés, tout

en n'améliorant pas la fusion. Des blocs de 32 x 32 pixels dégradent la qualité de la résolution.

Notons que les différents tests effectués sur des images de textures naturelles de Brodatz montrent qu'en général, l'utilisation de l'indice de similarité prenant en compte les différences entre fréquences d'occurrences est meilleur que celui qui utilise les attributs sur les fréquences d'apparition.

(a)

(b)

FIG. 3.43 - (a) Méthode basée sur les histogrammes d'ordre 2. Résultat de la segmentation par fusion de régions après un filtrage médian, avec une fenêtre de taille 16x16, et un seuil de fusion égal à 40% de la surface de la fenêtre. (b) Résultat de la segmentation par fusion de régions après un affinement par modélisation AR.

La segmentation à base des statistiques d'ordre deux est relativement robuste. On peut remarquer une meilleure limitation des frontières notamment pour la texture du sable (figure 3.42(a) en bas à gauche). La texture du bois (fond) est découpée en deux régions différentes ce qui n'est pas tout à fait incorrect. En effet, les étiquettes claires sont représentatives d'éventuels défauts de la texture du bois.

L'utilisation des statistiques d'ordre deux permet ainsi d'obtenir une segmentation plus stricte et moins sensible au bruit que la méthode fondée sur l'étude d'histogrammes locaux d'ordre un.

#### 3.3.1.2.4 Conclusion

Nous avons présenté une méthode de segmentation fondée sur une combinaison de deux processus de transformation de l'image. Une première transformation permet de pré-segmenter l'image en effectuant une classification automatique par multi-seuillage de ses

points après une analyse d'histogrammes locaux. L'objectif visé par cette opération est d'extraire des informations susceptibles de ne contenir que les éléments dont l'influence est la plus significative pour la caractérisation de différentes régions de l'image. Une seconde transformation consiste à segmenter l'image en minimisant un critère de similarité. Deux différents indices ont été étudiés. Le premier tient compte de la dynamique des histogrammes d'ordre un et des statistiques locales calculées à la fois sur l'image originale et celle des étiquettes résultantes de la pré-segmentation. Le second fait intervenir les histogrammes d'ordre deux. La fusion utilisant cet indice est plus robuste, elle a été expérimentée sur des images de textures naturelles. Elle donne des résultats satisfaisants, les différentes régions d'une image étant bien détectées au sens de la texture. Toutefois, malgré l'affinement effectué par la modélisation AR, la localisation des frontières des régions reste non correcte pour certaines configurations fines. L'intégration d'autres types d'informations et le recours à une coopération avec d'autres techniques sont indispensables pour optimiser les résultats. C'est l'objectif de la section suivante.

### 3.3.2 Interaction contour-régions uniformes et régions texturées

Afin de valider ou rejeter les contours non significatifs (situés à l'intérieur des zones texturées) issus de la coopération contours-régions du Module 1, nous avons intégré les résultats du second module.

Chaque contour situé à l'intérieur d'une région étiquetée comme texture est conservé ou enlevé en fonction de caractéristiques prenant en compte sa régularité, sa longueur, son épaisseur, son contraste. L'interaction de la détection de textures avec la procédure contours-régions est simple. A la fin de chaque itération, une correction du résultat de la détection des contours est effectuée en supprimant les courtes chaînes de contours non régulières et non fermées, se trouvant à l'intérieur d'une texture.

Pour estimer la régularité de la facette centrée sur le candidat à examiner, un indice de confiance a été défini, pour chaque point de contour, en fonction de la distribution des niveaux de gris de son voisinage local [172].

Soit  $I(i, j)$  le niveau de gris du point examiné. Le calcul de l'indice de confiance se fait de la manière suivante: on calcule les gradients interpolés  $G_i$  et leur moyenne  $\mu_{g_i}$

$$G_i(k) = P_{I_+}(k) - P_{I_-}(k) \quad \text{pour } k = \{-1, 0, 1\} \quad (3.39)$$

où  $P_{I_+}$  et  $P_{I_-}$  correspondent aux interpolations linéaires des points de l'image. Chaque  $P_I$  est déterminé par la combinaison linéaire de ses quatre voisins selon la figure 3.44:

$$P_I = (1 - a)(1 - b)I_{00} + (1 - a)bI_{01} + a(1 - b)I_{10} + abI_{11} \quad (3.40)$$

et

$$\mu_{g_i} = 1/3 \sum_{k=-1}^1 G_i(k) \quad (3.41)$$

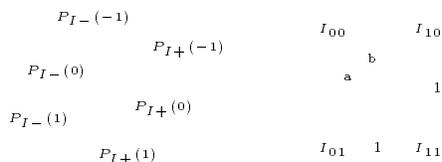


FIG. 3.44 - *Interpolation des points.*

L'inverse de la moyenne des écarts normalisée définit la régularité:  $Z = 1/Q$  où

$$Q = \frac{1/3 \sum_{k=-1}^1 |G_i(k) - \mu_{gi}|}{\mu_{gi}} \tag{3.42}$$

Pour une facette qui représente une variation homogène dans la direction du contour,  $Q$  tend vers zéro et par conséquent  $Z$  est grande. Par contre, si la présence d'un candidat est due à la texture, les fluctuations provoquées par celle-ci vont augmenter la valeur de  $Q$  et  $Z$  devient faible. L'indice de confiance du point candidat est donné par

$$C = \beta Z + \Delta[G_i(I(i, j))] \tag{3.43}$$

où  $\beta$  est un facteur correcteur et  $\Delta[.]$  est la valeur de la fonction SOC définie dans le paragraphe 3.3.1.1.1.

Dans la définition de  $C$ , lorsque la valeur de  $Z$  et/ou de  $\Delta[.]$  augmente, la valeur de  $C$  augmente aussi. Ceci est engendré pour une variation constante des voisins du point candidat et par le fait que la variation locale est importante pour les contours bien contrastés. Quand la valeur de  $\Delta[.]$  est très grande, le candidat doit être retenu même si  $Z$  est faible. C'est le rôle du paramètre  $\beta$ . Cette mesure doit préserver les détails évidents dans l'image originale.

### 3.3.3 Résultats expérimentaux

Les images de la figure 3.45 donnent des résultats expérimentaux pour l'image aérienne de la figure 3.22. Ces images présentent respectivement le résultat de multi-seuillage obtenu sur l'image Aérienne, son histogramme, l'histogramme global, l'histogramme global après la procédure de regroupement des masses ponctuelles, le résultat de la localisation de texture et le résultat de egmentation obtenus après intégration de la texture dans la coopération contours-régions. Dans cette expérimentation ainsi que dans l'ensemble des expérimentations suivantes, la taille des blocs est égale à 16. L'analyse est faite avec un chevauchement entre blocs de 8 pixels. L'introduction de la texture dans la coopération contours-régions uniformes permet d'obtenir des résultats de qualité correcte. Nous pouvons toutefois remarquer une

légère alteration des structures fines se situant entre des régions ayant des textures similaires.

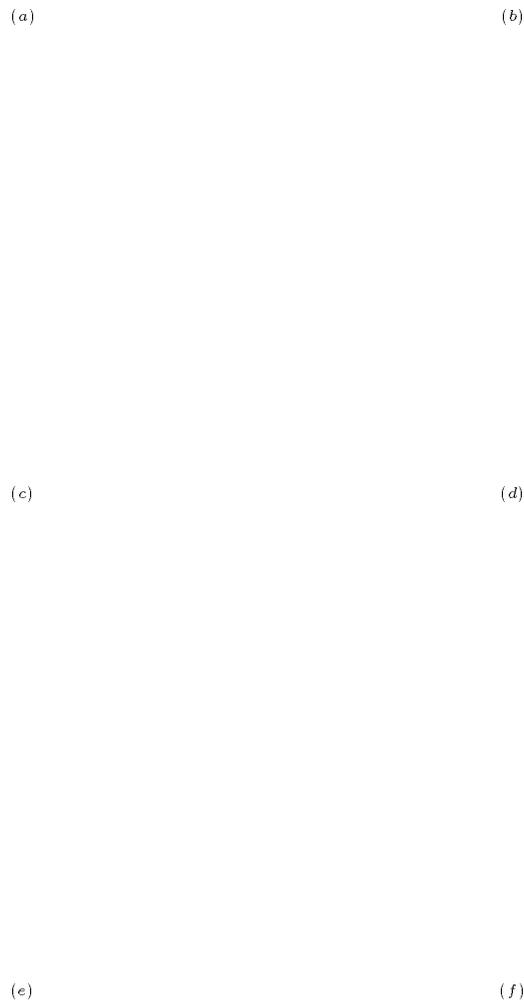


FIG. 3.45 - Segmentation obtenus après intégration de la texture dans la coopération contours-régions: l'image Aérienne. a) Son histogramme. b) Son histogramme global  $H_{glob}$ . c) Son histogramme global après regroupement des masses ponctuelles. d) Résultat du multi-seuillage (seuils: 98, 119, 135, 153, 176, 201). e) Résultat de la fusion des régions multi-seuillées. f) Segmentation après intégration de la texture.

### 3.3.4 Evaluation et étude comparative

Afin d'évaluer la méthode que nous avons développée et dans le but de comparer ses performances avec d'autres méthodes, nous reprenons l'image Paimpont (figure 3.46) prise par le satellite SPOT en mai 1987, car nous disposons de sa carte de vérité de terrain (figure 3.47) fournie par l'équipe COSTEL de l'Université de Rennes 2 (Climat et Occupation du Sol par TELédétection). Ceci facilite la comparaison des résultats expérimentaux des différentes méthodes. Comme la carte de vérité de terrain ne couvre pas entièrement l'image satellitaire, cette dernière a été masquée de sorte que toutes les deux aient le même support. Pour mettre en évidence le découpage cadastral, nous avons superposé à l'image originale les contours de parcelles dans la carte de vérité de terrain (figure 3.48).

La carte de vérité de terrain comprend 20 classes de végétation à identifier selon les travaux du COSTEL. Les numéros des classes et les statistiques telles que le nombre de points, la moyenne et l'écart-type de chaque classe sont données dans le tableau 3.12.

Pour faciliter la comparaison des résultats de classification, les 20 classes de végétation ont été fusionnées en 12 classes. Ce regroupement a été suggéré par les thématiciens du COSTEL. Nous avons renuméroté ces 12 classes par ordre croissant de la valeur moyenne des régions, et les statistiques correspondantes sont présentées dans le tableau 3.13. La carte de vérité de terrain des 12 classes est présentée en figure 3.47).

L'image de la figure 3.49(a) (resp. 3.49(b)) représente le résultat de la segmentation de l'image Paimpont obtenu par la méthode coopérative contours-régions (CR) (resp. la méthode coopérative contours-régions-textures (CRT)) développée superposée à l'image originale

Nous présentons dans les figures 3.50 et 3.51 les résultats après modélisation et classification en 12 classes des régions (constituées par les contours fermés de l'image 3.49(b)) obtenue par la méthode coopérative développée (CRT). La modélisation a été effectuée en utilisant respectivement les attributs de co-occurrences, les longueurs de plages et les deux premiers moments. Quant à la classification, elle a été réalisée en utilisant la méthode des K plus proches voisins.

Nous avons également traduit les résultats de classification en taux de bonne identification à l'aide de la carte de vérité de terrain 3.47. Le taux de bonne identification est le pourcentage des pixels de la m-ième classe dans la carte de vérité de terrain affectés à la n-ième classe de la segmentation finale. Les taux ainsi obtenus correspondant aux trois méthodes de modélisation sont donnés dans les tableaux 3.14 3.15, 3.16 3.17 et 3.18 3.19.

no vérité terrain	moyenne	écart-type	nbre de pixels	classe thématique
1	91,9	23,49	1706	Boisement lâche
2	120,3	21,75	280	Terre à nu (coupe rase)
3	118,4	23,52	957	Terre à nu (végétation)
5	76,9	18,61	5250	Futaie complète résineux
7	69,5	15,51	1160	Futaie claire résineux
8	86,7	18,71	1372	Futaie hétérogène conifères
10	117,2	26,93	3709	Taillis dégradé
11	144,3	20,62	6833	Taillis simple
12	95,4	21,62	2672	Réserve résineux
15	139,4	22,20	3193	Taillis sous futaie
16	115,7	34,98	791	Plantations conifères
17	137,9	30,70	2043	Gaulis conifères artificiel
18	117,6	17,70	1590	Bas perchis conifères artificiel
19	95,0	12,27	332	Haut perchis conifères artificiel
20	88,1	8,11	147	Gaulis conifères naturel
21	80,0	12,92	271	Perchis conifères naturel
22	107,2	14,59	22	Plantation feuillus
23	95,5	22,67	31	Plantation feuillus conifères
24	145,0	21,88	489	Haut perchis feuillus
25	139,7	14,09	648	Futaie feuillus

TAB. 3.12 - *Statistiques de 20 classes.*

n classe	n vérité de terrain	moyenne	écart-type	nbre de pixels
I	7	69,5	15,51	1160
II	5	76,9	18,61	5250
III	8	86,7	18,71	1372
IV	1	91,9	23,49	1706
V	12	95,4	21,62	2672
VI	18,19	113,7	18,93	1922
VII	2,3	118,9	23,14	1237
VIII	16,17,20,21,22,23	125,1	35,48	3305
IX	10,11	134,8	26,45	10542
X	15	139,4	22,20	3193
XI	25	139,7	14,09	648
XII	24	145,0	21,88	489

TAB. 3.13 - *Statistiques de 12 classes.*

FIG. 3.46 - *Image de la forêt de Paimpont masquée.*

FIG. 3.47 - *Carte vérité de terrain de la forêt de Paimpont VT. En contours (à gauche) et en régions (à droite).*

FIG. 3.48 - *Contours de la carte vérité de terrain de la forêt de Paimpont superposés à l'image originale.*

(a)

(b)

FIG. 3.49 - *Résultats de segmentation de l'image Paimpont obtenus par la méthode coopérative, (a) contours-régions (CR); (b) contours-régions-textures (CRT); développée superposée à l'image originale.*

(a)

(b)

FIG. 3.50 - *Résultat de segmentation de l'image Paimpont obtenu par la méthode CRT développée après classification en 12 classes en utilisant (a) les attributs de co-occurrence; (b) les attributs de longueurs de plages.*

FIG. 3.51 - *Résultat de segmentation de l'image Paimpont obtenu par la méthode CRT après classification en 12 classes en utilisant les 2 premiers moments.*

	1	2	3	4	5	6	7	8	9	10	11	12	total1
I	<b>94</b>	1	0	0	1	0	0	0	0	0	0	0	23222
II	2	30	7	11	29	5	2	0	0	1	0	5	1723
III	1	4	5	25	7	4	10	10	8	6	6	7	1269
IV	1	71	1	2	10	7	1	0	0	0	0	1	5523
V	0	<b>89</b>	0	4	3	2	0	0	0	0	0	0	1199
VI	4	42	3	32	4	1	2	2	1	0	1	0	1428
VII	2	3	3	4	7	3	5	9	<b>47</b>	1	1	9	11230
VIII	3	30	0	6	26	12	4	3	1	0	4	1	2763
IX	1	4	2	1	3	2	3	4	<b>41</b>	8	6	15	3255
X	1	9	3	3	15	5	4	3	13	14	4	19	3455
XI	0	2	0	18	9	9	14	8	6	0	14	13	1962
XII	0	0	0	0	6	0	0	8	24	0	0	<b>60</b>	683

TAB. 3.14 - Taux d'identification en (%) entre les classes de la vérité de terrain (première colonne) et les classes issues de la modélisation par les co-occurrences du résultat obtenu par la méthode CRT (première ligne).

	1	2	3	4	5	6	7	8	9	10	11	12
I	<b>94</b>	2	2	3	3	3	1	1	1	2	1	2
II	0	5	13	7	11	4	2	0	0	1	0	2
III	0	1	7	11	2	2	7	6	1	7	6	2
IV	0	<b>42</b>	4	5	13	17	4	1	0	0	0	2
V	0	11	0	2	1	1	0	0	0	0	0	0
VI	0	6	4	17	1	1	1	1	0	0	1	0
VII	1	4	31	15	19	14	29	50	<b>62</b>	13	11	25
VIII	0	9	1	6	16	16	7	4	0	1	10	1
IX	0	1	6	1	2	4	6	6	15	21	18	13
X	0	3	13	4	11	8	7	4	5	<b>43</b>	12	17
XI	0	0	1	12	4	8	16	7	1	0	<b>24</b>	6
XII	0	0	0	0	1	0	0	3	2	0	0	11
total2	23388	9378	920	2779	4423	2104	1812	2119	8600	1161	1141	3876

TAB. 3.15 - Taux d'identification en (%) entre les classes issues de la modélisation par les co-occurrences du résultat obtenu par la méthode CRT (première ligne) et les classes de la vérité de terrain (première colonne).

	1	2	3	4	5	6	7	8	9	10	11	12	total1
I	2	<b>92</b>	1	0	1	1	0	0	0	0	0	0	23222
II	0	2	5	13	38	5	8	6	6	7	3	1	1723
III	0	1	4	4	3	12	33	14	1	17	0	6	1269
IV	0	1	2	8	57	6	3	3	5	10	1	0	5523
V	0	0	1	0	<b>92</b>	2	0	0	0	4	0	0	1199
VI	0	4	3	22	28	3	2	11	3	7	4	7	1428
VII	0	1	2	9	42	5	12	13	2	4	4	2	11230
VIII	0	3	3	8	43	8	11	6	8	1	2	2	2763
IX	0	1	4	7	25	9	14	9	12	6	2	4	3255
X	0	1	6	3	13	6	22	10	14	13	4	1	3455
XI	0	0	2	6	8	10	12	31	6	0	10	6	1962
XII	0	0	0	8	19	1	6	0	0	<b>61</b>	0	4	683

TAB. 3.16 - Taux d'identification en (%) entre les classes de la vérité de terrain (première colonne) et les classes issues de la modélisation par les longueurs de plages du résultat obtenu par la méthode CRT (première ligne).

	1	2	3	4	5	6	7	8	9	10	11	12
I	<b>96</b>	94	15	3	1	6	1	2	2	2	2	1
II	0	0	5	6	4	3	3	2	4	4	4	1
III	0	0	3	2	0	5	9	4	1	7	0	8
IV	0	0	5	13	21	11	4	4	13	17	5	0
V	0	0	0	0	7	1	0	0	0	2	0	0
VI	0	0	2	9	3	1	1	4	2	3	4	10
VII	0	1	16	27	31	18	29	<b>34</b>	12	13	32	22
VIII	0	0	4	7	8	8	7	4	10	1	4	7
IX	0	0	8	6	5	11	10	7	16	6	6	16
X	0	0	12	3	3	7	16	8	<b>20</b>	15	11	3
XI	0	0	2	3	1	7	5	15	5	0	16	14
XII	0	0	0	2	1	0	1	0	0	13	0	3
total2	405	22595	1675	3506	15054	2766	4519	4172	2367	3143	1237	918

TAB. 3.17 - Taux d'identification en (%) entre les classes issues de la modélisation par les longueurs de plages du résultat obtenu par la méthode CRT (première ligne) et les classes de la vérité de terrain (première colonne).

	1	2	3	4	5	6	7	8	9	10	11	12	total1
I	94	0	0	0	0	0	1	0	0	0	0	0	23222
II	2	0	0	6	20	6	31	6	13	2	0	5	1723
III	1	0	0	0	3	1	9	16	28	16	13	7	1269
IV	1	0	0	0	60	3	19	7	2	1	0	2	5523
V	0	0	0	0	89	2	3	0	4	0	0	0	1199
VI	4	0	0	2	26	12	19	19	5	3	1	0	1428
VII	2	0	0	0	2	1	7	5	9	11	48	9	11230
VIII	3	0	0	1	4	0	52	12	11	7	2	2	2763
IX	1	0	1	1	3	2	1	3	4	10	42	17	3255
X	1	0	0	0	8	0	15	6	8	8	15	19	3455
XI	0	0	0	1	1	2	7	11	29	22	6	13	1962
XII	0	1	0	0	0	0	5	0	0	8	24	60	683

TAB. 3.18 - Taux d'identification en (%) entre les classes de la vérité de terrain (première colonne) et les classes issues de la modélisation par les deux premiers moments du résultat obtenu par la méthode CRT (première ligne).

	1	2	3	4	5	6	7	8	9	10	11	12
I	94	42	23	8	1	7	2	3	2	1	1	2
II	0	9	2	28	5	10	9	3	6	1	0	1
III	0	0	0	0	0	1	2	6	9	6	2	2
IV	0	0	1	5	47	17	18	13	3	1	0	0
V	0	0	0	0	15	2	1	0	1	0	0	0
VI	0	0	4	10	5	16	5	9	2	1	0	0
VII	1	0	4	2	4	10	13	17	27	36	61	16
VIII	0	0	0	7	1	1	24	11	7	5	0	0
IX	0	0	20	6	2	7	1	3	4	9	15	22
X	0	0	1	4	4	0	8	7	7	8	6	44
XI	0	0	0	6	0	4	2	7	15	13	1	0
XII	0	14	0	0	0	0	1	0	0	2	2	0
total2	23416	65	141	341	7012	1066	6057	3046	3953	3395	8821	1002

TAB. 3.19 - Taux d'identification en (%) entre les classes issues de la modélisation par deux premiers moments du résultat obtenu par la méthode CRT (première ligne) et les classes de la vérité de terrain (première colonne).

	$D_M(VT, RS_i)$	$D_M(RS_i, VT)$	Dissimilarité
Méthode CRT développée	0.004341	0.004411	0.004376
Méthode CR développée	0.004541	0.004617	0.004579

TAB. 3.20 - *Dissimilarités calculées entre les résultats obtenus par le système de segmentation (cas contour-région-texture CRT et cas contour-région CR) et la carte vérité de terrain VT.*

Le tableau 3.20 donne à titre indicatif, la dissimilarité normalisée entre la carte des frontières finales (figure 3.52(a)) (resp. (figure 3.52(b))) issue de la coopération contours-régions développée (méthode CR) (resp. coopération contours-régions-texture (méthode CRT)) et la vérité de terrain en contours  $VT$  de la figure 3.47.

(a)

(b)

FIG. 3.52 - *Contours du résultat de segmentation de l'image Paimpont obtenu par: (a) la méthode coopérative contours-régions (méthode CR); (b) la méthode coopérative contours-régions-textures (méthode CRT).*

Nous avons comparé les résultats de notre approche avec ceux obtenus par deux méthodes markoviennes : l'une par l'algorithme MAP-YOUNES et l'autre par l'algorithme SEM adaptatif MAP [43]. Ces deux algorithmes sont basés sur la modélisation des deux champs aléatoires : le premier champ, décrit par le modèle markovien "multilogistique" du premier ordre, correspond aux classes recherchées, et le deuxième champ, décrit par le modèle AR-2D gaussien, caractérise la texture des différentes classes. Les paramètres des modèles sont estimés par la procédure ICE ("Iterative Conditional Estimation") [267] pour l'algorithme MAP-YOUNES et par le SEM local [68] pour l'algorithme SEM adaptatif MAP. La classification des pixels s'est effectuée suivant le critère du Maximum A Posteriori. Les résultats

de ces deux algorithmes (superposés à la carte de vérité de terrain) sont donnés par la figure 3.53. Nous avons aussi calculé les taux de bonne classification correspondant à partir des tableaux de répartition des pixels de chaque classe. Les résultats sont présentés dans les tableaux 3.21 et 3.22.

(a)

(b)

FIG. 3.53 - *Résultat de segmentation de deux méthodes markoviennes. (a) algorithme MAP-YOUNES. (b) algorithme SEM adaptatif MAP.*

Bien que la segmentation apparaisse visuellement bonne, l'ensemble des résultats expérimentaux est médiocre en taux de bonne identification. Cela peut être dû à plusieurs facteurs :

- le décalage entre le moment de la prise de vue (image satellitaire acquise en mai 1987) et celui de la production du cadastre (carte de vérité de terrain réalisée en octobre 1989); Cela cause une mauvaise mise en correspondance des deux images (la plantation dans certaines zones a été déjà changée);
- le décalage entre les bords de zones dans l'image satellitaire et les tracés du cadastre; Ce défaut n'est pas négligeable. En effet, pour certaines zones, il peut provoquer une différence de surfaces correspondantes qui atteint jusqu'à 50%;
- la corrélation imparfaite entre les niveaux de gris de l'image satellitaire et la plantation réelle;

	1	2	3	4	5	6	7	8	9	10	11	12
I	1	65	16	9	3	5	0	0	0	0	0	0
II	3	36	26	17	9	6	3	0	0	0	0	0
III	4	10	26	17	15	13	10	2	2	0	0	0
IV	1	19	19	21	9	11	9	4	5	0	1	0
V	6	7	20	24	11	11	11	6	3	1	1	0
VI	0	0	1	5	6	19	26	27	8	3	3	1
VII	6	3	4	13	9	23	11	15	6	6	1	2
VIII	9	3	6	8	7	6	11	11	10	8	4	17
IX	4	1	2	4	3	6	11	10	8	13	20	18
X	0	0	1	2	3	4	6	12	12	19	23	17
XI	0	0	0	2	0	0	4	10	17	37	18	12
XII	2	0	0	4	2	2	9	10	16	5	13	37

TAB. 3.21 - Taux d'identification en (%) - algorithme MAP-YOUNES.

	1	2	3	4	5	6	7	8	9	10	11	12
I	8	64	12	7	4	5	0	0	0	0	0	0
II	8	38	20	14	9	8	3	0	0	0	0	0
III	4	17	21	17	12	18	7	4	1	0	0	0
IV	5	17	19	17	10	12	9	5	3	1	1	0
V	7	11	18	19	13	12	11	6	2	1	1	0
VI	0	1	2	4	7	21	27	26	6	3	3	2
VII	7	3	6	8	11	23	14	13	7	4	2	3
VIII	10	4	6	7	6	8	10	13	8	6	4	18
IX	4	1	2	3	3	7	11	11	8	12	20	19
X	0	0	1	2	3	4	8	12	14	12	21	23
XI	0	0	1	0	1	0	4	11	19	32	15	17
XII	0	0	0	0	3	4	7	13	10	9	10	39

TAB. 3.22 - Taux d'identification en (%) - algorithme SEM adaptatif MAP.

- la fragmentation des zones fortement texturées dans la carte des contours détectés;
- les défauts des modèles pour le regroupement de zones.

Afin de donner, à titre indicatif, une comparaison des résultats obtenus après modélisation des régions segmentées, nous avons considéré qu'un taux d'identification  $T_{iden}(m, n)$  supérieur à 40% est acceptable. Dans ce cas, nous avons constaté que la méthode de la matrice de cooccurrence et celle des deux premiers moments sont les plus performantes : plus de quatre importantes classes ont été identifiées avec un taux supérieur à 40% (tableaux 3.14 3.15 et 3.16 3.17). La méthode de longueurs de plages a identifiée trois classes (tableaux 3.18 3.19).

Nous avons également comparé les résultats de notre approche avec ceux des méthodes markoviennes [43]. D'après les tableaux 3.21 3.22, les deux algorithmes MAP-YOUNES et SEM adaptatif MAP donnent des résultats similaires : seule la classe I a été reconnue avec un taux de bonne identification supérieur à 40% (65% pour le premier et 64% pour le deuxième). En tenant compte du nombre des classes identifiées et des taux de bonne identification, notre approche est la plus efficace. Ceci peut également être constaté en comparant visuellement les résultats de segmentation.

## 3.4 Conclusion

L'approche coopérative développée permet d'obtenir une segmentation de l'image à la fois en primitives de type régions et de type contours. Elle se distingue des autres approches par l'introduction d'un mécanisme de correction et d'ajustement par le biais de boucles de retour et l'inclusion d'une mesure de cohérence des résultats intermédiaires, qui ont fait défaut à bon nombre d'approches coopératives.

Suivant la classification des différentes méthodes proposées dans la littérature, la coopération que nous proposons se situe dans le cadre des méthodes hybrides, relatives à la résolution de problèmes de la vision artificielle en utilisant des concepts de la "fusion de données" et de la "perception active". Beaucoup de parallèles peuvent être entrevus avec la coopération proposée par Bajcsy et al. Le système global est composé de plusieurs entités élémentaires. Chacune des composantes du système est dédiée à un traitement particulier. Le modèle de résolution du problème est une construction progressive de la solution. L'échec d'une des entités élémentaires du système ne doit pas remettre en cause le but global. Une introduction de mécanismes de correction et d'adaptations par le biais de boucle de retour est effectuée.

Toutefois, le système proposé comporte quelques différences notables qui se résument comme suit. Le but du système est l'exploration de données images dont on ne dispose pas de connaissances a priori. Le but global du système n'est pas défini explicitement pour une

application donnée. La coopération entre les principales composantes (détecteur de contours et extracteur de régions) de ce système est réalisée par une communication bidirectionnelle. Elle est réalisée à partir de sur-segmentations et dirigée dans une même direction jusqu'à la convergence vers la stabilité. Chaque composante produit un résultat approprié. Une évaluation de la cohérence des différents résultats intermédiaires est réalisée à la fin de chaque cycle à l'aide d'une dissimilarité prenant en compte à la fois la détection et la localisation des contours et des frontières des régions. La complexité algorithmique de cette mesure de dissimilarité dépend du nombre de points de contours et du nombre de points de frontières des régions dans l'image.

En tenant compte des défauts de la carte de vérité de terrain dont nous disposons, les résultats expérimentaux, exprimés par le pourcentage de bonne identification, montrent que la méthode de segmentation après modélisation avec les attributs de co-occurrences et les deux premiers moments sont les plus performantes car l'image traitée présente des zones uniformes avec des textures de nature micro-scopique. L'efficacité de notre approche de segmentation peut être constatée par la comparaison des résultats expérimentaux non seulement par rapport à l'image originale mais aussi par rapport aux résultats obtenus en utilisant les méthodes markoviennes.



## Conclusion et Perspectives

Isoler les différentes entités qui composent une image, constitue un des problèmes les plus importants en traitement et analyse d'images. Cette opération, quasi obligatoire dans tous les systèmes de vision artificielle, est la **segmentation**. Celle-ci conditionne fortement la qualité de l'interprétation, ce qui justifie les nombreux travaux qui lui ont été consacrés.

L'objectif de ce travail a été de concevoir un système de segmentation adaptable à l'analyse de plusieurs types d'images, riches en détails et contenant une grande variété de régions de nature uniforme ou texturée. Un intérêt tout particulier a été accordé à l'aspect automatique et non-supervisé du dispositif.

Nous nous sommes tout d'abord intéressés à l'étude de quelques concepts de fonctionnement du système visuel humain, puis, nous avons effectué un état de l'art des différents travaux réalisés dans le cadre de la segmentation d'images 2D fixes. Nous les avons classés en deux grandes catégories : **les approches non coopératives** et **les approches coopératives**. Dans la première catégorie, nous avons distingué trois types de méthodes : celles destinées à la détection des transitions locales (*contours*); celles adaptées à l'extraction des *régions uniformes* au sens des niveaux de gris; et celles dédiées à l'extraction des régions homogènes au sens de la *texture*. Les approches coopératives sont fondées sur la combinaison de plusieurs méthodes de la première catégorie, en particulier celles de détection de contours et d'extraction de régions uniformes. Dans cette catégorie, trois types de coopération ont été étudiés: la *coopération en "série"*, la *coopération "parallèle"* et la *coopération "hybride"*. De cette étude préliminaire, nous avons principalement dégagé les remarques suivantes:

Les méthodes de détection des primitives "*contours*" supposent généralement une connaissance a priori d'un modèle à détecter et opèrent d'une façon très localisée. Elles fournissent souvent des frontières de régions discontinues. De plus, elles sont sensibles aux variations spatiales brusques de l'image, qu'il s'agisse de bruit ou de texture, ce qui contribue à la production de faux contours à l'intérieur des régions.

Les méthodes d'extraction des primitives "*régions uniformes*" introduisent principalement des erreurs liées à la localisation des frontières des régions.

Les deux approches (contours et régions uniformes) présentent des caractéristiques qui sont complémentaires sur bien des points, en particulier la précision de la localisation des contours, la résistance au bruit et la fermeture naturelle des frontières des régions. Toutefois, en pratique, les représentations obtenues à l'aide de chacune de ces méthodes prise séparément ne vérifient que très rarement la complémentarité précédente.

Pour l'approche par extraction de primitives "*textures*", les résultats peu satisfaisants sont, essentiellement dus à deux facteurs: d'une part, à la difficulté d'extraction d'informations caractéristiques permettant d'augmenter la séparation entre les différents modèles de texture; d'autre part, à la fluctuation des niveaux de gris de zones homogènes.

L'ensemble de ces remarques montre bien qu'il est difficile d'aboutir à un résultat satisfaisant en appliquant une seule méthode. Aussi, pour espérer une segmentation fiable et suffisamment corrélée avec les entités qui composent la scène réelle, de nombreux travaux ont eu recours à la coopération de plusieurs des techniques précédentes.

L'automatisation de la tâche de segmentation et son adaptation à un large éventail de type d'image, a impliqué, d'une part, la multiplication des traitements, et d'autre part, l'adoption d'une démarche progressive où la formation des primitives s'opère de manière coopérative et guidée. Le système que nous avons mis au point entre donc dans le cadre des méthodes coopératives. Les éléments nouveaux dans cette coopération se résument, d'une part, par l'introduction d'un mécanisme de vérification de la cohérence par comparaison des résultats des méthodes qui coopèrent. Cet aspect fait défaut à un bon nombre d'approches coopératives. D'autre part, par l'inclusion de mécanisme de corrections et d'ajustement par le biais de boucle de retour.

L'architecture du système que nous proposons combine deux concepts. Le premier, fondé sur un procédé d'intégration d'informations issues de différentes méthodes, permet de tirer parti des avantages de chacune d'elles. Le second concept s'inspire de la perception active par l'introduction d'une boucle de retour dans le système. Aussi, il s'agit de réutiliser les acquis, mais aussi, d'améliorer la qualité des résultats de segmentation non supervisée, fondée sur la coopération entre les méthodes de détection de primitives de type "*contours*" et "*régions*" et de pousser l'investigation en introduisant de nouveaux mécanismes dans le processus de coopération. D'une part, pour permettre la remise en cause et la correction des résultats obtenus ainsi que l'ajustement des paramètres de contrôle des différentes techniques de segmentation. D'autre part, pour permettre la mesure de degré de cohérence des résultats de segmentation. Le système développé est composé de deux modules.

Le **premier** est dédié à l'extraction de régions uniformes ou faiblement texturées. Le principe est fondé sur une coopération itérative entre la détection des primitives "*contours*" et l'extraction des primitives "*régions uniformes*". Les deux méthodes de segmentation sont itérées avec des critères de plus en plus tolérants jusqu'à la conver-

gence vers des résultats compatibles et stables. Des affinements des résultats intermédiaires sont effectués. A la fin de chaque itération, la carte des régions est transformée en une carte de contours. Une mesure de similarité entre celle-ci et le résultat de la détection de contours est effectuée pour mesurer la cohérence des deux résultats. Cette mesure tient à la fois compte de la détection et de la localisation des contours. Le but est ainsi d'aboutir à une solution optimale au sens de la similarité entre les deux segmentations. La détection de contours s'effectue avec le filtre optimal de Canny-Deriche et l'extraction de régions uniformes s'appuie sur une approche d'agrégation de points.

Le **second module** fait intervenir les primitives “*textures*” afin de réactualiser et corriger les primitives contours et régions non significatives extraites au fur et à mesure des différentes itérations du premier module. La localisation des zones texturées dans l'image est fondée sur une combinaison de deux techniques que nous avons développées. La première s'appuie sur une classification automatique par multi-seuillage des niveaux de gris de l'image suite à une analyse d'histogrammes locaux transformés en utilisant des propriétés qui tiennent compte de la perception visuelle. La seconde technique s'appuie sur une procédure de fusion de régions minimisant un critère de similarité. Celui-ci prend en compte les histogrammes locaux d'ordre deux calculés sur l'image résultante de la première technique. L'algorithme de classification automatique par multi-seuillage a été testé sur une image synthétique (image de régions bruitées) et sur une image réelle (image de télédétection). Il a, par ailleurs, été comparé objectivement à d'autres techniques de multi-seuillage global, à l'aide d'une mesure similaire à la distance de Hausdorff moyennée. En plus des résultats satisfaisants obtenus par notre technique de multi-seuillage, elle est entièrement non-supervisée.

Le système combinant les deux modules a été appliqué sur plusieurs types d'images réelles et les résultats obtenus sont concluants. L'efficacité de notre approche s'est traduite, dans la plupart des cas examinés, par une détection cohérente des éléments représentatifs de l'image. Elle a été particulièrement constatée par les bons résultats expérimentaux obtenus sur les images de télédétection, et lors de la comparaison de ses résultats avec les segmentations obtenues en utilisant d'autres méthodes. Cette approche démontre ainsi l'importance de:

- l'intégration de primitives complémentaires (contours, régions, textures),
- l'interaction dans différentes directions,
- l'adoption d'une démarche progressive dans la formation des primitives,
- la validation de chaque primitive avec l'aide d'autres primitives et la mesure de sa cohérence, ce qui conduit à l'extraction de primitives fiables et pertinentes,
- et enfin, l'ajustement itératif des paramètres, qui permet leur adaptation à la spécificité de l'image analysée.

Les efforts futurs devront, d'une part, porter sur l'introduction d'une correction à partir du premier module vers le second (régions-uniformes vers textures) pour affiner les résultats de la localisation des frontières des textures. D'autre part, on peut envisager l'intégration d'autres méthodes tenant compte, de façon adaptative, de la résolution et des échelles des textures des images; ainsi que des critères ou méthodes permettant de prendre en considération l'aspect géométrique et structurel des primitives extraites au fur et à mesure des différentes itérations. De plus, une implantation de ce système sur une plateforme multi-tâches peut être envisagée afin de réaliser un réel parallélisme des différents traitements, que nous simulons actuellement de manière séquentielle.

## Annexe A

# Implémentation récursive de l'opérateur de Deriche

Dans cette annexe est présentée l'implémentation récursive de l'opérateur de Deriche (cf. 2.2.2.4.2)  $g(x) = Sxe^{-\alpha|x|}$  ainsi que son intégral  $q(x)$ .

### A.1 Cas monodimensionnel

Une mise en œuvre récursive du filtre optimal  $g(x)$  peut être obtenue en appliquant les techniques de la transformée en  $Z$ . Soient  $g(n)$  les échantillons de  $g(x)$  et  $G(Z)$  la transformée en  $Z$  correspondante :

$$G(Z) = \sum_{n=-\infty}^{\infty} g(n)Z^{-n} \quad (\text{A.1})$$

La séquence  $g(n)$  étant non causale, une étape préliminaire nécessaire est sa transformation en une somme de deux séquences causales. Pour cela, on réécrit  $g(n)$  sous la forme suivante d'une somme de deux séquences causales  $g_-(n)$  et  $g_+(n)$  telles que :

$$g(n) = g_-(n) + g_+(n) \quad (\text{A.2})$$

$$g_-(n) = \begin{cases} 0 & n \geq 0 \\ Sne^{\alpha n} & n < 0 \end{cases} \quad g_+(n) = \begin{cases} Sne^{-\alpha n} & n \geq 0 \\ 0 & n < 0 \end{cases}$$

En utilisant la transformée en  $Z$ , on obtient :

$$G(Z) = G_-(Z) + G_+(Z^{-1}) \quad (\text{A.3})$$

avec

$$G_+(Z^{-1}) = \frac{aZ^{-1}}{1 + b_1Z^{-1} + b_2Z^{-2}} \quad G_-(Z) = \frac{-aZ}{1 + b_1Z + b_2Z^2} \quad (\text{A.4})$$

avec

$$a = Se^{-\alpha}, \quad b_1 = -2e^{-\alpha}, \quad b_2 = e^{-2\alpha} \quad (\text{A.5})$$

$G_+(Z^{-1})$  (resp.  $G_-(Z)$ ) converge pour  $|e^{-\alpha}Z^{-1}| < 1$  (resp.  $|e^{\alpha}Z| > 1$ ). Comme toutes les singularités de  $G_+(Z^{-1})$  (resp.  $G_-(Z)$ ) sont à l'intérieur (resp. extérieur) du cercle unité pour  $\alpha$  réel positif, ces deux transformées en  $Z$  correspondent à deux fonctions de transfert de filtres récursifs stables et de second ordre. Le premier opérant de gauche à droite  $F_+$  et le second de la droite vers la gauche  $F_-$ .

En particulier, la réponse à  $x(m)$  comme entrée d'un système ayant  $g(n)$  comme réponse impulsionnelle peut être obtenue de manière récursive suivant le schéma suivant:

pour  $n = 1, \dots, M$ :

$$y_1(n) = x(n-1) + 2e^{-\alpha}y_1(n-1) - e^{-2\alpha}y_1(n-2) \quad (\text{A.6})$$

pour  $n = M, \dots, 1$ :

$$y_2(n) = x(n+1) + 2e^{-\alpha}y_2(n+1) - e^{-2\alpha}y_2(n+2) \quad (\text{A.7})$$

pour  $n = 1, \dots, M$ :

$$y(n) = ke^{-\alpha}[y_1(n) - y_2(n)] \quad (\text{A.8})$$

la constante  $S$  peut être fixée par la contrainte de normalisation suivante :

$$\sum_{n=0}^{+\infty} g(n) = - \sum_{n=-1}^0 g(n) = -1 \quad (\text{A.9})$$

ce qui donne :

$$S = -\frac{(1 - e^{-\alpha})^2}{e^{-\alpha}} \quad (\text{A.10})$$

L'importance de ces équations récursives de mise en œuvre tient au fait que le nombre d'opérations requis est extrêmement faible et surtout indépendant de la valeur de la résolution à laquelle les contours sont à détecter. En effet la forme du filtre déterminé par le paramètre  $\alpha$  peut varier mais le nombre d'opérations par point reste identique (5).

A titre de comparaison, une implémentation FIR de l'opérateur  $g(n)$  à l'aide de  $2N + 1$  coefficients non nuls aurait requis un nombre identique d'opérations par point de sortie. Un gain de  $(2N + 1)/5$  en résulte en faveur de la mise en œuvre récursive qui présente en plus l'avantage de ne pas introduire de bruit de troncature. A titre d'exemple une implémentation sur 16 bits de cet opérateur de détection sous forme RIF pour  $\alpha = 0.5$  et  $\alpha = 0.25$  aurait nécessité respectivement 57 et 105 opérations par point.

En appliquant la même technique, on peut aisément montré que les échantillons  $q(n)$  de la primitive  $q(x)$  de  $g(x)$  possèdent aussi la bonne propriété de pouvoir être mis en œuvre de manière récursive. Ce résultat sera utilisé pour l'extension au cas 2D. En effet :

$$q(n) = k(\alpha|n| + 1)e^{-\alpha|n|} \quad (\text{A.11})$$

Montrons que  $q(n)$  peut aussi être implémenté de manière récursive. Soit  $H(Z)$  sa transformée en  $Z$

$$H(Z) = H_-(Z) + H_+(Z^{-1}) \quad (\text{A.12})$$

avec

$$H_+(Z^{-1}) = \frac{a_0 + a_1Z^{-1}}{1 + b_1Z^{-1} + b_2Z^{-2}} \quad H_-(Z) = \frac{a_2Z + a_3Z^2}{1 + b_1Z + b_2Z^2} \quad (\text{A.13})$$

avec

$$a_0 = k, \quad a_1 = k(\alpha - 1)e^{-\alpha}, \quad a_2 = a_1 - kb_1, \quad a_3 = -kb_2 \quad (\text{A.14})$$

Ayant le même dénominateur que  $G_+(Z^{-1})$  (resp.  $G_-(Z)$ ), ces deux transformées en  $Z$  correspondent de la même manière à deux filtres récursifs stables et de second ordre. Le premier opérant de gauche à droite  $F+$  et le second de la droite vers la gauche  $F-$ .

La sortie  $y(m)$  résultat de la convolution de la séquence  $x(m)$  avec la réponse impulsionnelle  $q(n)$  s'obtient alors comme suit:

pour  $n = 1, \dots, M$ :

$$y_1(n) = k[x(n) + e^{-\alpha}(\alpha - 1)x(n - 1)] + 2e^{-\alpha}y_1(n - 1) - e^{-2\alpha}y_1(n - 2) \quad (\text{A.15})$$

pour  $n = M, \dots, 1$ :

$$y_2(n) = k[e^{-\alpha}(\alpha + 1)x(n + 1) - e^{-2\alpha}x(n + 2)] + 2e^{-\alpha}y_2(n + 1) - e^{-2\alpha}y_2(n + 2) \quad (\text{A.16})$$

pour  $n = 1, \dots, M$ :

$$y(n) = y_1(n) + y_2(n) \quad (\text{A.17})$$

La constante  $k$  est fixée par la contrainte de normalisation suivante :

$$\sum_{n=-\infty}^{+\infty} S(n) = 1 \quad (\text{A.18})$$

On obtient alors :

$$k = \frac{(1 - e^{-\alpha})^2}{1 + 2\alpha e^{-\alpha} - e^{-2\alpha}} \quad (\text{A.19})$$

Cette étape de lissage nécessite seulement 8 opérations par point indépendamment de la résolution  $\alpha$  utilisée. Ceci doit être comparé aux  $2N + 1$  opérations requises par un filtre RIF de  $2N + 1$  coefficients. A titre d'exemples pour  $\alpha = .5$  et  $.25$ , une implementation RIF sur 16 bits de ce filtre de lissage aurait demandé respectivement 57 et 113 opérations par point.

## A.2 Cas bidimensionnel

Dans cette partie, la procédure de généralisation au cas 2D de l'opérateur de détection de Deriche [108] est présentée. Pour cela, on crée pour la direction en  $x$  (resp.  $y$ ) un masque 2D noté  $SS_x$  (resp.  $SS_y$ ) produit du détecteur aligné en  $x$  (resp.  $y$ ) avec un opérateur de projection parallèle à la direction en  $y$  (resp.  $x$ ). Une implémentation totalement récursive en résulte si on choisit comme opérateur de lissage l'opérateur  $q(x)$  présenté dans la section précédente.

$$SS_x(m, n) = k_1 m e^{-\alpha \cdot |m|} \cdot k_2 (\alpha |n| + 1) e^{-\alpha \cdot |n|} \quad (\text{A.20})$$

$$SS_y(m, n) = k_2 (\alpha |m| + 1) e^{-\alpha \cdot |m|} \cdot k_1 n e^{-\alpha \cdot |n|} \quad (\text{A.21})$$

On note  $SS$  l'opérateur de lissage 2D correspondant à ces deux gradients directionnels :

$$SS(m, n) = k_2 (\alpha |m| + 1) e^{-\alpha \cdot |m|} k_2 (\alpha |n| + 1) e^{-\alpha \cdot |n|} \quad (\text{A.22})$$

La figure A.1 illustre les opérateurs de lissage et de dérivation et permet de comparer leur forme à l'opérateur Gaussien ainsi qu'à sa première dérivée.

FIG. A.1 - Les opérateurs de lissage et de dérivation. C1: exponentiel; C2: gaussienne

Ces trois masques séparables peuvent être efficacement mis en œuvre de manière récursive selon les étapes suivantes:

Dans une première phase, un filtrage récursif est appliqué à chaque ligne de l'image  $x(m, n)$  à traiter comme suit:

pour  $n = 1, \dots, N$  et  $m = 1, \dots, M$ :

$$y_1(m, n) = a_1 x(m, n) + a_2 x(m, n-1) + b_1 y_1(m, n-1) + b_2 y_1(m, n-2) \quad (\text{A.23})$$

pour  $n = N, \dots, 1$  et pour  $m = 1, \dots, M$ :

$$y_2(m, n) = a_3 x(m, n+1) + a_4 x(m, n+2) + b_1 y_2(m, n+1) + b_2 y_2(m, n+2) \quad (\text{A.24})$$

pour  $n = 1, \dots, N$  et pour  $m = 1, \dots, M$ :

$$r(m, n) = c_1 (y_1(m, n) + y_2(m, n)) \quad (\text{A.25})$$

Une seconde phase applique alors au résultat  $r(m, n)$  le second filtre au niveau de chaque colonne pour obtenir le résultat final  $y(m, n)$ .

pour  $m = 1, \dots, M$  et pour  $n = 1, \dots, N$ :

$$y_1(m, n) = a_5 r(m, n) + a_6 r(m - 1, n) + b_1 y_1(m - 1, n) + b_2 y_1(m - 2, n) \quad (\text{A.26})$$

pour  $m = M, \dots, 1$  et pour  $n = 1, \dots, N$ :

$$y_2(m, n) = a_7 r(m + 1, n) + a_8 r(m + 2, n) + b_1 y_2(m + 1, n) + b_2 y_2(m + 2, n) \quad (\text{A.27})$$

pour  $n = 1, \dots, N$  et pour  $m = 1, \dots, M$ :

$$y(m, n) = c_2 (y_1(m, n) + y_2(m, n)) \quad (\text{A.28})$$

Dans [108], il est montré que cette structure permet de mettre en œuvre efficacement les opérations de lissage et de différentiation directionelles. Ainsi, une convolution avec l'opérateur à deux indices donné par 90, correspondant à un lissage de l'image originale, peut être effectuée en choisissant le jeu de paramètres suivant :

$$a_1 = a_5 = k; a_2 = a_6 = ke^{-\alpha}(\alpha - 1); a_3 = a_7 = ke^{-\alpha}(\alpha + 1); a_4 = a_8 = -ke^{-2\alpha}; c_1 = c_2 = 1; \quad (\text{A.29})$$

Une dérivée directionelle en  $x$  peut être calculée par le jeu de paramètres suivant :

$$a_1 = 0; a_2 = 1; a_3 = -1; a_4 = 0; c_1 = -(1 - e^{-\alpha})^2; \quad (\text{A.30})$$

$$a_5 = k; a_6 = ke^{-\alpha}(\alpha - 1); a_7 = ke^{-\alpha}(\alpha + 1); a_8 = -ke^{-2\alpha}; c_2 = 1; \quad (\text{A.31})$$

La dérivée directionelle en  $y$  peut être calculée en interchangeant les paramètres  $a_i$  avec  $a_{i+4}$  et  $c_1$  avec  $c_2$ .

Pour toutes ces applications, les coefficients  $b_1$ ,  $b_2$  et  $k$  sont donnés par:

$$b_1 = 2e^{-\alpha}; b_2 = -e^{-2\alpha}; k = \frac{(1 - e^{-\alpha})^2}{1 + 2\alpha e^{-\alpha} - e^{-2\alpha}} \quad (\text{A.32})$$

Cette mise en œuvre permet la convolution de l'image en seulement 13 opérations par point de sortie indépendamment de la taille du filtre spécifié par le paramètre  $\alpha$ .



## Annexe B

# Lissage adaptatif utilisant la diffusion anisotropique

Dans le processus d'analyse d'image, l'étape de pré-traitement est généralement nécessaire pour régulariser les données et atténuer le bruit qui s'introduit durant les étapes d'acquisition de l'image. Elle a également pour but d'améliorer l'homogénéité des régions tout en conservant, voire en renforçant, le contraste entre régions adjacentes.

Il existe principalement deux catégories de filtres de lissage: les opérateurs non-linéaires et les opérateurs linéaires. Un filtre est non-linéaire quand les pixels d'une image ne contribuent pas de la même façon au lissage. Généralement, cela dépend de la position spatiale et de la valeur du niveau de gris du pixel. A l'inverse, un filtre est linéaire quand la contribution de tous les points de l'image au lissage est la même. Dans le cas linéaire, le lissage peut être réalisé par des opérateurs spatiaux de type moyenneur ou gaussien. Comme tous les points de l'image sont traités de la même manière, les contours ayant un gradient faible risquent d'être perdus. Afin d'éviter les distorsions subies par ce type de contours une restauration et un lissage adaptatif sont préférables [25] [320].

Dans cette annexe est présentée une approche adaptative de lissage basée sur la diffusion anisotropique.

Ce type d'opérateur est adaptatif car les paramètres sont calculés selon le contexte de l'image. Conçu pour préserver au mieux les contours tout en atténuant les hautes fréquences, ce lissage consiste à convoluer l'image de façon itérative par un masque dont les coefficients de pondération correspondent à une mesure de degré de continuité du signal image en chaque point. Dans [309] il est montré que ce processus itératif engendre une implémentation approximative discrète de l'équation de diffusion anisotropique de la chaleur telle que celle proposée par Perona et Malik [274] en résolvant l'équation

$$\frac{\partial I}{\partial t} = \Delta(c\Delta I) \quad (\text{B.1})$$

où  $I$  correspond à l'image originale,  $c$  est le coefficient de diffusion,  $t$  représente le temps et  $\Delta$  l'opérateur gradient.

Soient  $I^{(k)}(x, y)$  et  $\Delta^{(k)}(x, y)$  l'intensité et le module du gradient du pixel  $(x, y)$  examiné à l'itération  $(k)$ . La procédure de lissage itératif est régie par une relation de la forme:

$$I^{(k+1)}(x, y) = \frac{\sum_{i=-1}^1 \sum_{j=-1}^1 I^{(k)}(x+i, y+j) \omega^{(k)}(x+i, y+j)}{\sum_{i=-1}^1 \sum_{j=-1}^1 \omega^{(k)}(x+i, y+j)} \quad (\text{B.2})$$

où

$$\omega^{(k)}(x, y) = e^{-\frac{(\Delta^{(k)}(x, y))^2}{2s^2}} \quad (\text{B.3})$$

$s$  est un facteur d'échelle.

L'image lissée à l'itération  $(k+1)$  correspond à l'image lissée à l'itération  $(k)$  pondérée par une fonction  $w(\cdot)$  approchant l'inverse de l'exponentielle du module du gradient au pixel examiné.  $w(\cdot)$  est proche de zero au niveau des discontinuités (gradient élevé) et proche de 1 dans les zones homogènes. A chaque itération seules les régions homogènes sont prises en compte dans le processus de lissage.



# Abbreviations bibliographie

## Congrès

CV	"Handbook of Pattern Recognition and Image Processing: Computer Vision"
CVPR	"IEEE Computer Society Conference on Computer Vision and Pattern Recognition"
ECCV	"European Conference on Computer Vision"
GRETSI	"Conférence-Colloque sur le Traitement du Signal et des Images"
ICIP	"International Conference on Image Processing"
ICPR	"International Conference on Pattern Recognition"
ICSP	"International Conference on Signal Processing"
RFIA	"Reconnaissance des Formes et Intelligence Artificielle"
SPIE	"Society of Photo-Optical Instrumentation Engineers"

## Journaux

AI	"Artificial Intelligence"
CS	"Cybernetics and Systems"
CGA	"IEEE Computer Graphics and Applications"
CVGIP	"Computer Vision Graphics and Image Processing"
GMIP	"Graphical Models and Image Processing"
IJCV	"International Journal of Computer Vision"
IJPRAI	"International Journal of Pattern Recognition and Artificial Intelligence"
IJRR	"International Journal of Robotics Research"
IPL	"Information Processing Letters"
IS	"Information Sciences"
ISA	"Information Sciences Applications"
IU	"Image Understanding"
IVC	"Image and Vision Computing"
JCSS	"Journal of Computer and Systems Sciences"
JMIV	"Journal of Mathematical Imaging and Vision"
JPDC	"Journal of Parallel and Distributed Computing"
JVCIR	"Journal of Visual Communication and Image Representation"
MVA	"Machine Vision and Applications"
P-IEEE	"Proceedings of the IEEE"
PR	"Pattern Recognition"
PRL	"Pattern Recognition Letters SIAM"
JC	"SIAM Journal on Computing"
SP	"Signal Processing"
SV	"Spatial Vision"
TCS	"Theoretical Computer Science"
T-IP	"IEEE Transactions on Image Processing"
T-IS	"IEEE Transactions on Information and Systems"
TOG	"ACM Transactions on Graphics"
T-PAMI	"IEEE Transactions on Pattern Analysis and Machine Intelligence"
T-PDS	"IEEE Transactions on Parallel and Distributed Systems"
T-RA	"IEEE Transactions on Robotics and Automation"
T-SMC	"IEEE Transactions on Systems, Man, and Cybernetics"
TS	"Revue Traitement du Signal"
VC	"The Visual Computer"

# Bibliographie

- [1] A.S. ABUTALEB, Automatic thresholding of gray-level pictures using two dimensional entropy, CVGIP 47 pp. 22-32,1989.
- [2] F. ADE, Characterization of textures by eigenfilters, SP 5 pp. 451-457,1983.
- [3] M. ABBASI-DEZFOULI and T.G. FREEMAN, Patch matching in stereo-images based on shape. In Spatial Information from Digital Photogrammetry and Computer Vision, pp. 1-8, Munich, sep. 1994.
- [4] R. AZENCOTT and C. GRAFFIGNE, Non supervised segmentation using multi-Level Markov Random Fields. Proc. ICPR, pp. 201-205, The Hague Nd. Sep. 1992.
- [5] J. ALOIMONOS, Purposive and qualitative active vision. In ICPR, pp. 346-360, Atlantic City, New Jersey, 1990.
- [6] J. ALOIMONOS, What I Have Learned. CVGIP: IU, 60(1) pp. 74-85, July 1994.
- [7] A. ACKAH-MIEZAN and A. GAGALOWICZ, Discrete models for energy minimizing segmentation. ICCV, pp. 200-207, Berlin, Mai 1993.
- [8] J. K. AGGARWAL and N. NANDHARKUMAR, Multisensor fusion for automatic scene interpretation. In R. C. Jain and A. K. Jain, editors, Analysis and Interpretation of Range Images, pp. 339-361. Springer-Verlag.
- [9] H. ANDERSON, Edge-detection for object recognition in aerial photographs. Tec. Rep. MS-CIS-87-96, University of Pennsylvania, Grasp Laboratory, 1987.
- [10] I.E. ABDOU and W.K. PRATT, Quantitative design and evaluation of enhancement/ thresholding edges detectors, P-IEEE 76 pp. 753-763,1979.
- [11] N. AHUJA and A. ROSENFELD, A note on the use of second-order gray-level statistics for threshold selection, T-SMC, SCM-8, pp. 895-899, 1978.
- [12] N. AHUJA and A. ROSENFELD, Mosaic models for textures, T-PAMI-3(1) pp. 1-11,1981.
- [13] N. AHUJA and M. TUCERYAN, Extractions of Early Perceptual Structure in Dot Patterns: Integrating Region, Boundary and Component Gestalt. PR, vol. 48, pp. 304-356, 1989.
- [14] J. ALOIMONOS, I. WEISS, and A. BANDYOPADHYAY, Active Vision. In Proc. of Int. Conf. Computer Vision, 1987.
- [15] R. AZENCOTT, Champs markoviens et imagerie. Congrès Reconnaissance des Formes et Intelligence Artificielle, pp. 1183-1191, 1987.
- [16] S. BARONTI, A. CASINI, F. LOTTI, L. FAVARO, and V. ROBERTO, Variable Pyramid Structures for Image Segmentation, CVGIP (49), No. 3, pp. 346-356, March 1990.
- [17] D. BOURENNANE et al, Généralisation du filtre de Canny-Deriche pour la détection des contours de formes quelconques, 14ième colloque GRETSI, Juan-les-Pins, pp. 715-718, Sept. 1993.
- [18] A.J. BADDELEY, An error metric for binary images. Robust Computer Vision, Ruwiedel (ed.); Karlsruhe, Wichmann pp. 59-78; 1992.
- [19] R. BAJCSY, Active Perception. Proceedings of The IEEE, 76(8) pp. 996-1005, August 1988.
- [20] D.H. BALLARD, Generalizing the Hough transform to detect arbitrary sharps, PR 13 pp. 111-122,1981.

- [21] D.H. BALLARD, Reference frames for animate vision. In Proc. Int. Joint Conf. on AI, volume 2, Detroit, Michigan, pp. 1635-1641, August 1989.
- [22] D.H. BALLARD and C.M. BROWN, *Computer Vision*, Prentice-Hall HJ 1982.
- [23] J. BENOIS and D. BARBA, Image segmentation by region-contour cooperation as basis for efficient coding scheme, Proc. of the ICPR, The Hague Nd. Sep. 1992.
- [24] A.C. BOVIK, M. CLARK and W.S. GEISLER, Multichannel texture analysis using localized spatial filters. T-PAMI-12(1) pp. 55-73,1990.
- [25] L. BEAUREPAIRE, Filtrage d'images numériques. Thèse de doctorat, Université de Rennes I, 1996.
- [26] F. BERGHOLM, Edge focusing, T-PAMI-9(6) pp. 726-741, 1987.
- [27] J. BESAG, Spatial interaction and statistical analysis of lattice system, J. Roy. Stat. Soc. Series B 36 pp. 192-226,1974.
- [28] J. BESAG, On the Statistical analysis of dirty pictures, volume 48 of B. J. Royal Stat. Soc., 1986.
- [29] J.C. BEZDEK, Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York, 1981.
- [30] B. BHANU and O. FAUGERAS, Segmentation of images having unimodal distributions, T-PAMI 4, pp. 408-419, 1982.
- [31] C.R. BRICE and C.L. FENNEMA, Scene analysis using Regions. AI, 1 pp. 205-226, 1970.
- [32] S. BASU and K.S. FU, Image segmentation by syntactic method, PR 20(1) pp. 33-39, 1987.
- [33] S.T. BARNARD and M.A. FISCHLER, Computational stereo. Comput. Surveys, (4) pp. 553-572, dec 1982.
- [34] H. BUNKE, H. FEISTL, H. NEIMANN, G. SAGERER, F. WOLF, and G. Zhou, Smoothing, thresholding, and contour extraction in images from gated blood pool studies, in First IEEE Symp. on Medical Imaging and Image Interpretation, Berlin 1982.
- [35] O. BAUJARD et C. GARBAY, KISS: Un système de vision multi-agents. Actes du 8eme congrès de Reconnaissance des Formes et Intelligence Artificielle, pp. 89-98, Lyon-Villeurbanne, Novembre 1991.
- [36] J.R. BEVERIDGE, J. GRIFFITH, R.R. KOHLER, A.R. HANSON and E.M. RISEMAN, Segmenting images using localized histograms and region merging. IJCV, 2 pp. 311-347, Janvier 1989.
- [37] P. BURT, T.H. HONG, and A. ROSENFELD, Segmentation and estimation of image region properties, through cooperative hierarchical computation, T-SMC-11(12), 1981.
- [38] P.J. BESL, and R.C. JAIN, Segmentation through variable-order surface fitting, PAMI(10), No. 2, pp. 167-192, March 1988.
- [39] R. BAJCSY and L. LIEBERMAN, Texture gradient as depth cue, CGIP 5 pp. 52-67, 1976.
- [40] C. BOUMAN and B. LIU, Multiple resolution segmentation of images, T-PAMI-13(2) pp. 99-113,1991.
- [41] S.S. BLACKMAN, Theoretical approaches to data association and fusion. In SPIE Sensor Fusion, volume 931, 1988.
- [42] A. BLAKE, Comparison of the deterministic and stochastic algorithms for visual reconstruction, T-PAMI-11(1) pp. 2-12, 1989.
- [43] J.M. BOUCHER et P. LENA, Evaluation et comparaison de méthodes de segmentation en imagerie SPOT, rapport du projet ECOSS, Collaboration Bretagne Image, ENST de Bretagne, Janvier 1992.
- [44] R. BAJCSY, M. MINTZ, and E. LIEBMAN, A common frame work for edge and region growing, Report Univ. of Pens. Feb. 1986.
- [45] D. BARBA and J. RONSIN, New method in texture analysis in the context of image segmentation, 2nd European SP Conference, pp. 283-286, Germany, 1983.
- [46] A.D. BRINK, Gray-level thresholding of images using a correlation criterion, PRL 9 pp. 335-341, 1989.

- [47] P. BRODATZ, Textures- A Photographic album for artists and designers, Dover NY 1966.
- [48] C.M. BROWN, Toward general vision. CVGIP: IU, 60(1) pp. 89-91, July 1994.
- [49] S. BOUKHAROUBA, J.M. REBORDAO, and P.L. WENDEL, An amplitude segmentation method based on the distribution function of an image, CVGIP. 29, pp. 47-59, 1985.
- [50] C. BOUMAN and M. SHAPIRO, A multiscale random field model for bayesian image segmentation. T-IP, 3(2) pp. 162-177, Mars 1994.
- [51] N. BOUJEMAA, G. STAMON, and A. GAGALOWICZ. Modélisation floue pour la segmentation d'images. In RFIA, pp. 163-173, Paris, France, 1994.
- [52] F. BELLET, M. SALOTTI, C. GARBAY, Une approche opportuniste et coopérative pour la vision bas niveau, TS, Vol.12(5) pp. 479-493, 1995.
- [53] A. BLAKE and A. ZISSERMAN, Visual reconstruction. MIT Press, Cambridge, USA, August 1987.
- [54] P. BONNIN et B. ZAVIDOVIQUE, La segmentation coopérative: Comment combiner détection de contours et croissance de régions? 14ième colloque GRETSI, Juan-les-Pins, pp. 755-758, Sept. 1993.
- [55] H. CHABBI, Construction de facettes 3D par stéréovision. Thèse de Doctorat de l'INPL, février 1993.
- [56] B. CHALMOND, Image restoration using an estimated markov model. SP, 15 pp. 115-129, 1988.
- [57] B. CHANDA, On Image enhancement and threshold selection using the gray level cooccurrence matrix, PRL 3(4) pp. 243-251, 1985.
- [58] C.C. CHU and J.K. AGGARWAL, The Integration of image segmentation maps using region and edge information, T-PAMI(15), No. 12, pp. 1241-12, December 1993.
- [59] A. CASALS, J. AMAT and A. GRAU, Texture parametrization method for image segmentation. ECCV, volume 588, pp. 160-164, Italy 1992.
- [60] J. CANNY, A computational approach to edge detection, T-PAMI-8(6) pp. 679-698, November 1986.
- [61] C. CARIOU, Analyse spectrale d'images numériques - application à la télédétection spatiale des océans, Thèse de Doctorat de l'Université de Bretagne Occidentale, 1991.
- [62] P.B. CHOU and C.M. BROWN, The theory and practice of bayesian image labeling. IJCV, 4(3) pp. 185-210, 1990.
- [63] J.L. CROWLEY and J.M. BEDRUNE, Integration and control of reactive visual process. In Proc. ECCV, volume 2, pp. 47-58, Stokholm, Sweden, May 1994.
- [64] K. CHEHDI et D. COQUIN, Binarisation de plusieurs classes d'images par seuillage local optimal maximisant un critère d'homogénéité, 13ème colloque GRETSI, Juan-les-pins, Septembre 1991.
- [65] F.S. COHEN and D.B. COOPER, Simple parallel ierarchical and relaxation algorithms for segmenting noncausal Markovian Random Fields. T-PAMI 9(2) pp. 195-219, Mars 1987.
- [66] K. CHEHDI, C. CARIOU and C. KERMAD, Image segmentation and texture classification using local thresholds and 2D AR modelling, in Proc. of EUSIPCO, Edinburgh, pp. 30-33, Sept. 1994.
- [67] C.I. CHANG, K. CHEN, J.W. WANG, and M.L.G. ALTHOUSE, A relative entropy-based approach to image thresholding, PR(27), No. 9, pp. 1275-1289, September 1994.
- [68] G. CELEUX and J. DIEBOLT, The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. Computation Statistics Quaterly, 2 pp. 73-82, 1985.
- [69] J.P. COCQUEREZ et J. DEVARS, Détection de contours dans les images aériennes: nouveaux opérateurs. TS Vol 2, Numero 1, 1985.
- [70] J.M. CHASSERY et Y. ELOMARY, Coopération contours actifs et multi-résolution en segmentation d'images, 15ième colloque GRETSI, Juan-les-pins, Septembre 1995.

- [71] C. CSEKE and I.Z. FAZEKAS, Comments on gray level thresholding of images using correlation criterion, PRL 11 pp. 709-711, 1990.
- [72] J.M. CHASSERY and C. GARBAY, An iterative segmentation method based on a contextual color and shape criterion, T-PAMI-6 pp. 794-800, 1984.
- [73] R.W. CONNERS and C.A. HARLOW, A theoretical comparison of texture algorithms, T-PAMI-2(3) pp. 204-222, 1980.
- [74] C.C. CHEN and C.L. HUANG, Markov random fields for texture classification. PRL 14 pp. 907-914, Novembre 1993.
- [75] A. CHEHIKIAN, Binarisation d'images: deux solutions à ce problème, TS, Vol 6, N1, 1989.
- [76] A. CHEHIKIAN, Filtres récursifs pour l'estimation du gradient et la détection de contours par interpolation splines, TS, Vol 14, N1, 1997.
- [77] H. CAILLOL, A. HILLION, and W. PIECZYNSKI. Segmentation contextuelle non supervisée d'images utilisant une modélisation statistique floue. In RFIA, pp. 145-151, Paris, France, 1994.
- [78] T. CAELLI and B. JULESZ, On perceptual analyzers underlying visual texture discrimination. Biological Cybernetics, 28(29) pp. 201-214, 1978.
- [79] G.R. COSS and A.K. JAIN, Markov random field for texture models, T-PAMI-5 pp. 25-39, 1983.
- [80] C.K. CHOW and T. KANEKO, Boundary detection of radiographic images by a threshold method, in Proceedings, IFIP Congress, pp. 130-134, 1971.
- [81] C.K. CHOW and T. KANEKO, Automatic boundary detection of left ventricle from cineangiograms, Comput. Biomed. Res. 5, pp. 338-410, 1972.
- [82] R. CHELLAPPA and R.L. KASHYAP, Texture synthesis using 2D non-causal AR models, IEEE ASSP-33 pp. 194-203, 1985.
- [83] K. CHEHDI et Q.M. LIAO, Détection de contours basée sur la perception visuelle en vue de la segmentation d'images, 14ième colloque GRETSI, Juan-les-Pins, pp. 739-742, Sept. 1993.
- [84] D. CANO and T.H. MINTH, Texture synthesis using hierarchical linear transformation, SP 15 pp. 131-148, 1988.
- [85] J.M. CHASSERY and M. MELKEMI, A segmentation method using Voronoi diagrams in a split and merge environment. Progress in Image Analysis and Processing, pp. 44-48, Positano, Septembre 1989.
- [86] J.M. CHASSERY et A. MONTANVERT, Géométrie discrète en analyse d'images. Eds. Hermès, Paris 1991.
- [87] H. I. CHRISTENSEN and C. B. MADSEN, Purposive reconstruction: a reply to a computational and evolutionary perspective on the role of representation in vision. CVGIP: IU, 60(1) pp. 86-88, July 1994.
- [88] R.C.K. CHUNG and R. NEVATIA, Recovering building structures from stereo. IEEE CVPR'92, pp. 64-73, 1992.
- [89] Y.Q. CHEN, M.S. NIXON and D.W. THOMAS, Statistical geometrical features for texture classification. PR, 28(4) pp. 537-549, 1995.
- [90] T.N. CORNSTWEET, Visual Perception, Academic Press, N.Y. 1970.
- [91] H. COSTIN. On object classification by means of fuzzy sets'theory. In ICPR, pp. 458-461, La Haye, 1992.
- [92] P.C. CHEN and T. PAVLIDIS, Segmentation by texture using a cooccurrence matrix and a split and merge algorithm. CVGIP, 10 pp. 172-182, 1979.
- [93] J.P. COCQUEREZ et S. PHILIPP, Analyse d'images: filtrage et segmentation. Collection Enseignement de la physique. Masson, Paris, 1995.
- [94] J.P. COCQUEREZ et S. PHILIPP, Comparaison de méthodes de segmentation d'images, conf. invité, 15ème GRETSI, Juan-les-Pins, sept. 1995.
- [95] B. CHARROUX, S. PHILIPP, J.P. COCQUEREZ, Système de vision mettant en oeuvre une coopération d'opérateurs de segmentation guidée par l'interprétation, 10ème congrès AFCET RFIA, Rennes, 1996.

- [96] J.L. CROWLEY and J.M. BEDRUNE. Integration and control of reactive visual process. In Proc. ECCV, volume 2, pp. 47-58, Stokholm, Sweden, May 1994.
- [97] M.G. COX, J.M. REHG and S. HINGORANI, A Bayesian multiple-hypothesis approach to edge grouping and contour segmentation. IJCV, vol. 11, no 1, pp. 5-24, 1993.
- [98] X. CHEN and F. SCHMITT, Split and merge image segmentation based on Delaunay triangulation. The 7th Scandinavian Conference on Image Analysis (SCIA, IAPR), pp. 910-917, Azlborg, Dannemark, Août 1991.
- [99] A. CHU, S.M. SEHGAL, and J.F. GREENLEAF, Use of gray level distribution of run lengths for texture analysis, PRL 11 pp. 415-420,1990.
- [100] CVGIP: Image Understanding, 60(1) pp. 65-118, July 1994.
- [101] S. CASTAN, J. ZHAO et J. SHEN, Une famille de détecteur de contours basée sur le filtre exponentiel optimal, Actes du Cong. RFIA Paris 1989.
- [102] U.R. DHOND and J.K. AGGARWAL, Structure from stereo - a review. IEEE Transactions on Systems, Man, and Cybernetics, (6) pp. 1489-1510, dec 1989.
- [103] A.L. DAVIS, A survey of edge detection techniques, CGIP 4 pp. 251-259,1976.
- [104] M. DERRAS, M. BERDUCAT, C. DEBAIN and P. BONTON, Unsupervised regions segmentation: real time control of an upkeep machine of natural spaces. European Conference on Computer Vision, volume B, pp. 207-212, Mai 1994.
- [105] L.S. DAVIS, M. CLEARMAN, and J.K. AGGARWAL, An empirical evaluation of generalized cooccurrence matrices, T-PAMI-3(2) pp. 214-221,1981.
- [106] H. DERIN and H. ELLIOT, Modeling and segmentation of noisy and textured images using Gibbs random fields. T-PAMI, 9(1) pp. 39-55, Janvier 1987.
- [107] H. DERIN, H. ELLIOT, R. CRISTI and D. GEMAN, Bayes smoothing algorithms for segmentation of binary images modelled by MRF, T-PAMI-9 pp. 707-720,1987.
- [108] R. DERICHE, Using Canny's criteria to derive an optimal edge detector recursively implemented, IJCV 1(2) pp. 167-187,1987.
- [109] R. DERICHE, J.P. COCQUEREZ, and G. ALMOUZNI. An efficient method to built early image description. In 9th ICPR, Rome, Italie, nov 1988.
- [110] R. DERICHE, Fast algorithms for low-level vision T-PAMI-18(1) pp. 679-683,1990.
- [111] R.O. DUDA and P.E. HURT, Pattern Classification and Scene Analysis, Wiley NY 1973.
- [112] D. DUNN, W.E. HIGGINS and J. WAKELEY, Texture segmentation using 2-D Gabor elementary functions. T-PAMI 16(2) pp. 130-149, Février 1994.
- [113] E. DIDAY, J. LEMAIRE, J. POUJET et F. TESTU, Eléments d'analyse de données, Eds. Dunod, Paris, 1982.
- [114] R.C. DUBES and A.K. JAIN, Random field models in image analysis. Journal of Applied Statistics, 16(2) pp. 131-164, 1989.
- [115] A.P. DEMPSTER, N.M. LAIRD and D.B. RUBIN, Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc., 39(1) pp. 138, série B, Methodological, 1977.
- [116] V. DI GESU and M. C. MACCARONE, Features selection and possibility theory. PR, (1) pp. 63-72, 1986.
- [117] F. DERAVID and S.K. PAL, Gray level thresholding using second-order statistics, PR Lett. 1, pp. 417-422, 1983.
- [118] D. DUBOIS et H. PRADE, Théorie des possibilités. Masson, 1985.
- [119] J. DOLAN and E. RISEMAN, Computing curvilinear structure by token-based grouping. In CVPR, Urbana Champaign, (USA), pp. 264-270. juin 1992.
- [120] P. De SOUZA, Edge detection using sliding statistical tests, CVGIP 23 pp. 1-14,1983.

- [121] O. FAUGERAS, Artificial 3D Vision, MIT Press, Cambridge 1993.
- [122] W. FREI and C.C. CHEN, Fast boundary detection: a generalization and a new algorithm, IEEE C-26(10) pp. 988-998, 1977.
- [123] J. FERBER et M. GHALLAB Problématique des univers multi-agents intelligents, Actes des journées nationales PRC IA, Toulouse, France, pp. 295-320, 1988.
- [124] G. FEKETE, J.O. EKLUNDH, and A. ROSENFELD, Relaxation: evaluation and applications, IEEE Trans. Pattern Anal. Mach. Intell., PAMI-3, pp. 460-469, 1981.
- [125] W.D. FISHER, On grouping for maximum homogeneity, JASA, vol 53, pp. 789-798, 1958.
- [126] M.M. FLECK. Spectre: an improved phantom edge finder. In Proc. 5th Alvey Vision Conference, pp. 127-132, 1989.
- [127] S.M.X. FERNANDO and D.M. MONRO, Variable thresholding applied to angiography, in Proceedings, 6th International Conference on PR, 1982.
- [128] S.K. FU and J.K. MU, A survey on image segmentation PR 13, pp. 3-16, 1981.
- [129] J.M. FRANCOS, A.Z. MEIRI and B. PORAT, A unified texture model based on 2D Word-Like decomposition, T-PAMI-41(8) pp. 665-678, 1993
- [130] C. FERNANDEZ-MALOIGNE, A. SMOLARZ et H. PAVIE, Comparaison d'algorithmes de sélection de seuils à l'aide des matrices de cooccurrence et de longueurs de plage. APII26, pp. 437-452, 1992.
- [131] G. FLOUZAT, J. SERRA, et Y. MERGHOUB, Aide automatique la compréhension des images de télédétection: modélisation des caractères morphologiques de la texture. 1er Colloque Image, Biarritz Mai 1984, pp. 735-740.
- [132] K.S. FU, Syntactic image modelling using stochastique tree grammars, CGIP 12 pp. 136-152, 1980.
- [133] A. GAGALOWICZ, Vers un modèle de texture. Thèse, Université Pierre et Marie Currie, PARIS VI, 6 Mai 1983.
- [134] M.M. GALLOWAY, Texture analysis using gray level run lengths. CVIP, 4 pp. 172-179, 1975.
- [135] A. GAGALOWICZ et O. MONGA, Un algorithme de segmentation hiérarchique, Actes du 5ième congrès RFIA, Grenoble Nov. 1985.
- [136] J.P. GOMBOTTO, A new approach to combining region growing and edge detection, PRL 14 pp. 869-875, 1993.
- [137] L.V. GOOL, P. DEWAELE and A. OOSTERLINCK, Survey: Texture analysis anno 1983. CVGIP, 29 pp. 336-357, 1985.
- [138] S. GEMAN and D. GEMAN, Stochastic relaxation, Gibbs distribution, and the bayesian restoration of images. T-PAMI, 6(6) pp. 721-741, Novembre 1984.
- [139] D. GEIGER and F. GIROSI, Parallel and deterministic algorithms for MRFs: surface reconstruction and integration, Proc. of ECCV, Antibes 1990, pp. 89-98.
- [140] D. GEMAN, S. GEMAN, C. GRAFFIGNE and P. DONG, Boundary detection by constrained optimization. T-PAMI 12(7) pp. 609-628, Juillet 1990.
- [141] W.I. GROSKY and R. JAIN, Optimal quadtree for image segmentation, T-PAMI-5 pp. 77-83, 1983.
- [142] C.C. GOTLIEB and H.E. KREYSZIG, Texture discriptors based on cooccurrence matrices, CVGIP 51 pp. 70-86, 1990.
- [143] A. GRUEN, O. KUEBLER, and P. AGOURIS (eds.). Automatic extraction of man-made objects from aerial and space images, Monte Verita Ascona (Switzerland), Birkhauser, Apr. 1995.
- [144] T.D. GARVEY, J. D. LOWRANCE, and M. A FISCHLER, An inference technique for integrating knowledge from disparate sources. In Proc. of the Int. Joint Conf. on AI, 1981.
- [145] A. GAGALOWICZ and S.De MA, Synthesis of natural textures, in Proc. of 6th ICPR, pp. 1081-1086, Munich 1982.
- [146] G. GIDEON and G. MEDIONI, Perceptual grouping using global saliency-enhancing operators. In Proc. ICPR 92, pp. 99-103, La Haye, 1992.

- [147] S. GROSSBERG and E. MINGOLLA, Neural dynamics of perceptual grouping: textures, boundaries and emergent segmentations. *Perception and Psychophysics*, 38-2 pp. 141-171, 1985.
- [148] R.C. GONZALEZ and P. WINTZ, *Digital Image Processing*. Addison-Wesley publishing compagny, 1992. Seconde edition.
- [149] D. GEIGER and A. YUILLE, A common framework for image segmentation, *IJCV* 6(3) pp. 227-243,1991.
- [150] J.P. GOMBOTTO and O. MONGA, A parallel and hierarchical algorithm for region growing. In *ICPR*, San Francisco June 1985.
- [151] W. HOFF and N. AHUJA, Surfaces from stereo: integrating feature matching, disparity estimation, and contour detection. *T-PAMI*, (2) pp. 121-136, feb 1989.
- [152] J.P. HATON, Les systèmes à base de connaissances en reconnaissance et en interprétation de formes, *COGNITIVA*, Paris, La Vilette, pp. 73-80, 18-22 May 1987.
- [153] R.M. HARALICK. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5) pp. 786-803, 1979.
- [154] R.M. HARALICK, Digital step edge from zero crossing of second directional derivatives, *T-PAMI*-6(1) pp. 58-68, January 1984.
- [155] R.M. HARALICK, Statistical approaches to textures, in *Handbook of PR and Image Processing*, T.S.Young and K.S.Fu Ed. Academic Press NY 1986.
- [156] R.M. HARALICK, Image analysis using mathematical morphology, *T-PAMI*-9 pp. 532-550,1987.
- [157] J.F. HADDON and J.F. BOYCE, Image segmentation by Unifying Region and Boundary Information, *SPIE* 933, 1988.
- [158] J.F. HADDON and J.F. BOYCE, Image segmentation unifying region and boundary information, *T-PAMI*-12(10) pp. 929-948,1990.
- [159] T.C. HENDERSON and O. D. FAUGERAS, High level multisensor integration. In *SPIE*, *Sensor Fusion: Spatial Reasoning and Scene Interpretation*, volume 1003, pp. 307- 314, 1988.
- [160] S. HOUZELLE et G. GIRAUDON, Segmentation en régions par modélisation de matrices de cooccurrence. 8ième congrès RFIA, pp. 1177-1183, Lyon, Nov. 1991.
- [161] D.P. HUTTENLOOCHER, G. KLANDERMAN and J. RUCKLIDGE, Comparating images using the Hausdorff distance, *T-PAMI* Vol 15, Numero 9, pp. 850-863, September 1993.
- [162] A. HUERTAS G. MEDIONI, Detection of intensity changes with subpixel accuracy using laplacien-gaussian masks, *T-PAMI* Vol 8, Numero 5, pp. 651-664, September 1986.
- [163] A. HILLION, P. MASSON et C. ROUX, Une méthode de classification de textures par extraction linéaire non paramétrique de caractéristiques, *TS* 5(4) pp. 273-279, 1988.
- [164] B.K.P. HORN, *Robot Vision*, McGrawHill, NY, 1986.
- [165] S.L. HOROWITZ and T. PAVLIDIS, Picture segmentation by a directed split and merge procedure. *Proc. ICPR*, volume 424-433, 1974.
- [166] S.L. HOROWITZ and T.PAVLIDIS, Picture segmentation by a tree transversal algorithm, *JACM* 23 pp. 368-388,1976.
- [167] R.M. HARALICK and L.G. SHAPIRO, Survey, image segmentation techniques, *CVGIP*, Vol 29, pp. 100-132, 1985.
- [168] R.M. HARALICK, K. SHANMUGAM and I. DINSTEIN, Textural features for image classification. *T-SMC*, 3(6) pp. 610-621, Novembre 1973.
- [169] R. HORAUD and T. SKORDAS, Stereo correspondence through feature grouping and maximal cliques. *T-PAMI*, (11) pp. 1168-1180, nov 1989.
- [170] C.A. HARLOW, M.M. TRIVEDI and R.W. CONNERS, Use of texture operators in segmentation. *Optical Engineering*, 25(11) pp. 1200-1206, Novembre 1986.
- [171] M.H. HUECKEL, A local visual edge operator which recognizes edges and lines, *JACM* 20 pp. 634-647,1973.

- [172] R.M. HARALICK and L.T. WATSON, A facet model for image data, *CGIP* 15 pp. 113-126,1981.
- [173] R. HUMMEL and S. ZUCKER, On the foundations of relaxation labeling processes. *IEEE trans. on Pattern Analysis and Machine Intel ligen*ce, 5(3), 1983.
- [174] E. ITO and J. ALOIMONOS, Determining transformation parameters from images: theory. In *Proc. IEEE Conference on Robotics and Automation*, 1987.
- [175] A.K. JAIN, *Digital Image Processing*, Wiley Interscience Publication, NY 1989.
- [176] R. JAIN, Expansive vision. *CVGIP: IU*, 60(1) pp. 86-88, July 1994.
- [177] G. JOHANNSEN and J. BILLE, A threshold selection method using information measures, in *Proceedings, 6th ICPR*, Munich, Germany, pp. 140-143, 1982.
- [178] A.K. JAIN and F. FARROKHANIA, Unsupervised texture segmentation using Gabor filters, *PR* 24 pp. 1167-1186,1991.
- [179] B. JULESZ, Textons, the elements of texture perception, and their interactions. *Nature*, 290 pp. 91-97, 1981.
- [180] B. JULESZ, Experiments in visual perception of texture, *Sci. Amer.* 232(4) pp. 2-11,1975.
- [181] T. KANADE, Survey region segmentation: signal vs semantics. *CVIP*, 13 pp. 279-297, 1980.
- [182] R. KARA-FALAH et P. BOLON, Mesure de dissimilarité entre deux segmentations, 14ième colloque GRETSI, pp. 763-766, Septembre 1993.
- [183] R. KARA-FALAH et P. BOLON, Une technique d'intégration des résultats en segmentation d'images, 15ième colloque GRETSI, pp. 573-576, Juan les pins, Septembre 1995.
- [184] R. KARA-FALAH, P. BOLON, and J.P. COCQUEREZ, A region-region and region-edge cooperative approach of image segmentation, *Proc. ICIP*, Vol.3 pp. 470-474, 1994.
- [185] J.M. KELLER, S. CHEN and R.M. CROWNOVER, Texture description and segmentation through fractal geometry. *CVGIP*, 45 pp. 150-166, 1989.
- [186] C. KERMAD, La Segmentation d'images 2D fixe: une revue. Rapport de Recherche interne LASTI-IMAGE 07/94, Juil. 1994.
- [187] C. KERMAD, K. CHEHDI et C. CARIOU, Segmentation d'images par multi-seuillage d'histogramme et fusion de régions minimisant un critère de similarité, *Proc. 15ième colloque GRETSI*, pp. 641-644, Juan-les-Pins Sept. 1995.
- [188] C. KERMAD, K. CHEHDI and C. CARIOU, Image segmentation by an iterative region-contour control minimizing a convergence criterion, *Proc. IEEE ICSP'96*, Pekin, Oct. 1996.
- [189] R.L. KASHYAP, R. CHELLAPPA, and A. KHOTANZAD, Texture classification using features derived from random field models, *PRL* 1 pp. 43-50,1982.
- [190] C. KERVANN et F. HEITZ, Segmentation non supervisée des images naturelles texturées: Une approche statistique. *Traitement du Signal*, 11(1) pp. 31-41, 1994.
- [191] J. KITTLER and J. ILLINGWORTH, Minimum error thresholding, *PR* 19, pp. ; 41-47, 1986.
- [192] J. KITTLER and J. ILLINGWORTH, On threshold selection using clustering criterion, *T-SMC-15*, pp. 652-655, 1985.
- [193] R. KIRSCH, Computer determination of the constituent structure of biological images, *Computers and Biological Research* 4(3) pp. 315-328,1971.
- [194] M. KUNT and R. LEONARDI, Adaptive region growing technique using polynomial function for image approximation, *SP* 11(1),1986.
- [195] R. KOHLER, A segmentation system based on thresholding, *Comput. Graphics Image Process.* 15, pp. 319-338, 1981.
- [196] V. KOIVUNEN and M. PIETIKAINEN, Combined edge and region based method for range image segmentation, *Proc. of SPIE Int. robotics and computer vision conf. IX*, pp. 501-512, Boston Nov. 1990.

- [197] H. KIM and R. PARK, Extraction of periodicity vectors from structural textures using projection information, PRL 11(9) pp. 625-630, 1990.
- [198] R.L. KIRBY and A. ROSENFELD, A note on use of (gray level, local average gray level) space as an aid in thresholding selection, T-SMC-9, pp. 860-864, 1979.
- [199] E. KROTKOV, Exploratory visual sensing for determining spatial layout with an agile stereo camera system. Tec. Rep. MS-CIS87-29, University of Pennsylvania, April 1987.
- [200] J.N. KAPUR, P.K. SAHOO, and A.K.C. WONG A new method for gray-level picture thresholding using the entropy of histogram, CVGIP. 29, pp. 273-285, 1985.
- [201] M. KASS, A. WITKIN, and D. TERZOPOULOS, Snakes: Active contour models, IJCV 1 pp. 321-331, 1988.
- [202] H.Y. KIM and H.S. YANG, A Systematic Way for Region-Based Image Segmentation Based on Markov Random Field Model. PRL, 15 pp. 969-976, 1994.
- [203] Z. KATO, J. ZERUBIA and M. BERTHOD, Unsupervised parallel image classification using a hierarchical markovian model. Rapport de recherche 2528, INRIA, Avril 1995.
- [204] Y.H. KATZ, Pattern recognition of meteorological satellite cloud photography, Proceedings, Third Symp. on Remote Sensing of Environment, pp. 173-214, 1965.
- [205] K.I. LAWS, Textures energy measures. In Proc. of IU Workshop, pp. 47-51, 1979.
- [206] S.U. LEE, S.Y. CHUNG, and R.H. PARK, A comparative performance study of several global thresholding techniques for segmentation, CVGIP 52 pp. 171-190, 1990.
- [207] S. LAKSHMANAN and H. DERIN, Simultaneous parameter estimation and segmentation of Gibbs Random Fields using simulated annealing. T-PAMI, 11(8) pp. 799-813, Août 1989.
- [208] J.G. LEU, Detecting the spatial structure of textures based on shape analysis, CGIP 31 pp. 67-88, 1985.
- [209] S.Y. LU and K.S. FU, A Syntactic approach to texture Analysis. CVGIP, 7 pp. 303-330, 1978.
- [210] Q.M. LIAO, Detection de contours et segmentation d'images, application à la télédétection et à la biologie marine. Thèse de doctorat, Université de Rennes I, 1995.
- [211] R.C. LUO and M.G. KAY, Multisensor integration and fusion in intelligent systems. T-SMC, (5) pp. 901-931, September 1989.
- [212] Y.W. LIM and S.U. LEE, On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. PR, (9) pp. 935-952, 1990.
- [213] R.C. LUO and M.H. LIN, Robot multi-sensor fusion and integration pp. Optimum estimation of fused sensor data. In Proc. of the Int. Conf. on Robotics and Automation, pp. 1076-1081, April 1988.
- [214] M.D. LEVINE and A.M. NAZIF, An optimal set of image segmentation rules. PR, 2 pp. 243-248, 1984.
- [215] M.D. LEVINE and A.M. NAZIF, Dynamic measurement of computer generated image segmentation. T-PAMI 7(2) pp. 155-164, Mars 1985.
- [216] D.G. LOWE, Perceptual organization and visual recognition. Kluwer Academic Publishers, Hingham, 1985.
- [217] H. LUIJENDIJK, Automatic threshold selection using histograms based on the count of 4-connection regions, PRL 12(4) pp. 219-228, 1991.
- [218] Y. LECHEVALIER et J. LEVY VEHEL, Arthur: un système d'analyse de texture. 8e congrès RFIA, volume 3, pp. 1185-1193, Lyon, Villeurbanne, Nov. 1991.
- [219] A. MITICHE and J.K. AGGARWAL, Image segmentation by conventional and information integration techniques. IVC 3(2) pp. 50-60, 1985.
- [220] R. MA, Analyse de séquences d'images stéréoscopiques en vue de la détection d'obstacles dans un routier. Thèse de doctorat, Université de Nice - Sophia-Antipolis, 1994.

- [221] S.G. MALLAT, A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. T-PAMI 11(7) pp. 674-693, Juillet 1989.
- [222] D. MARR, Vision. Eds. W.H. Freeman and Company, New York 1982.
- [223] J.L. MARROQUIN, Optimal Bayesian estimation for image segmentation and surface reconstruction, MIT-AI-MEMO 893, 1985.
- [224] A. MARTELLI, An application heuristic search methods to edge and contours detection, ACM 19 pp. 73-83,1976.
- [225] G. MATHERON, Éléments pour une théorie des milieux poreux, Masson, Paris 1967.
- [226] F. MANGIN, M. BERTHOD, and J. ZERUBIA, A cooperative network for contour grouping. In Proc. ICPR, The Hague, Netherlands, sep 1992.
- [227] B.S. MANJUNATH and R. CHELLAPPA, Unsupervised texture segmentation using markov random field models. T-PAMI, 13(5) pp. 478-482, Mai 1991.
- [228] M. MELKEMI and J.M. CHASSERY, Edge-region segmentation processing based on generalized voronoï diagram representation, Proc. of the ICPR, The Hague Nd. Sep. 1992.
- [229] J.W. MODESTINO, R.W. FREI and A.L. VICKERS, Stochastique models generated by random tessellations of the plane, CGIP 12(1) pp. 74-98,1980.
- [230] D.L. MILGRAM and D.J. HERMAN, Clustering edge values for threshold selection, Comput. Graphics Image Processing, Vol 10, pp272-280, 1979.
- [231] D. MARR and E. HILDRETH, Theory of edge detection, Proc. of the Royal Society of London, Series B 207 pp. 187-217,1980.
- [232] B.H. McCORMICK and S.N. JAYAMUMURTHY, Time series model for texture synthesis, Int. J. Comput. Inform. Sci. 3(4) pp. 329-343.1974.
- [233] D. MASSON, I.J. LAUDER, D. RUTRITZ, and G. SPOWART, Measurement of C-Bands in human chromosomes, Comput. Biol. Med. 5, pp . 179-201, 1975.
- [234] T. MATSUYAMA, S. MIURA and M. NAGAO, Structural analysis of natural textures by fourier transformation. CVGIP, 24 pp. 347-362, 1983.
- [235] J.L. MARROQUIN, S. MITTER and T. POGGIO, Probalistic solution of three ill-posed problems in computational vision, J of the Ame. Stat. Assoc. 82 pp. 76-89,1987.
- [236] G. MEDIONI and R. NEVATIA, Segment-based stereo matching. CVGIP Graphical Models and Image Processing, 31 pp. 2-18, 1985.
- [237] O. MONGA, An optimal region growing algorithm for image segmentation. International Journal of PR and AI, 1(3 et 4) pp. 351-375, 1987.
- [238] T.H. MORRIN, A black-white representation of gray-scale picture, IEEE Trans. Comput. 23, pp. 184-186, 1974.
- [239] A. MURTHY and S.K. PAL, Fuzzy thresholding: mathematical framework, bound functions and weighted moving average technique. PRL , (11) pp. 197-206, 1990.
- [240] N. METROPOLIS, A. ROSENBLUTH, M. ROSENBLUTH, H. TELLER and E. TELLER, Equation of state calculation by fast computing machines. Journal Chemical Physics, 21 pp. 1087-1091, Juin 1953.
- [241] A. MURTHY and S. K. PAL, Fuzzy thresholding: mathematical framework, bound functions and weighted moving average technique. PRL , (11) pp. 197-206, 1990.
- [242] R. MUZZOLINI, Y.H. YANG and R. PIERSON, Texture characterization using robust statistics. PR, 27(1) pp. 119-134, 1994.
- [243] J.W. MODESTINO and J. ZHANG, A Markov random field modelBased approach to image interpretation. T-PAMI 14(6) pp. 606-615, Juin 1992.

- [244] D.M. MCKEOWN and Y.C. HSIEH, Hierarchical waveform matching: A new feature-based stereo technique. IEEE CVPR, pp. 513-519, 1992.
- [245] R. NELSON and J. ALOIMONOS, Obstacle avoidance: towards qualitative vision. In Proc. ICCV, December 1988.
- [246] V.S. NALWA and T.O. BINFORD, On detecting edges, IEEE Trans. Pattern Analysis and Machine, PAMI-8(6) pp. 699-714, 1986.
- [247] A. NAZIF and M.D. LEVINE, Low level segmentation: An expert system, T-PAMI-6(5),1984.
- [248] M. NAGAO and M. MATSUYAMA, Edge preserving smoothing. In CVGIP, volume 9, pp. 394-407, 1979.
- [249] V.S. NALWA and E. PAUCHON, Edge aggregation and edge description. CVGIP, vol. 40, pp. 79-94, 1987.
- [250] Y. NAKAGAWA and A. ROSENFELD, Some experiments on variable thresholding, PR 11, pp. 191-204, 1979.
- [251] R. OHLANDER, K. PRICE and D.R. REDDY, Picture segmentation using recursive region splitting method. CGIP, 8 pp. 313-333, 1978.
- [252] P.P. OHANIAN and R.C.DUBES. Performance evaluation for four classes of textural features. PR, 25(8) pp. 819-933, 1992.
- [253] Y. OHTA, K. TAKANO, and K. IKEDA, A high speed stereo matching system based on dynamic programming. In ICCV, pp. 335-342, London, jun 1987.
- [254] N. OSTU, Threshold selection method from gray-level histogram, T-SMC-8, pp. 62-66, 1978.
- [255] T. PONG and al. Experiment in image segmentation using a facet model grower, CGIP 25 pp. 1-12,1984.
- [256] D.P. PANDA, Segmentation of FLIR images by pixel classification, TR-508, University of Maryland, Computer Science Center, 1977.
- [257] T. PAVLIDIS, Structural pattern recognition, Springer-Verlag, New York, 1977.
- [258] T. PAVLIDIS, Algorithms for Graphical and Image Processing, Springer-Verlag, 1982.
- [259] T. PAVLIDIS, Why progress in machine vision is so slow, PRL 13 pp. 221-225, 1992.
- [260] N.R. PAL and D. BHANDARI, Image thresholding: some new techniques, SP 33 pp. 139-158, 1993.
- [261] S. PELEG, A new probabilistic relaxation scheme, T-PAMI-2, pp. 362-369, 1980.
- [262] A. PENTLAND, Fractale Textures, in Proc. of the IJCAI, Germany, Aug. 1983.
- [263] A.P. PENTLAND, Fractal based description of natural scenes. T-PAMI, 6 pp. 661-674, Novembre 1984.
- [264] W.K. PRATT, O. FAUGERAS and A. GAGALOWICZ, Visual discrimination of stochastique texture fields, T-SMC-8 pp. 796-804,1978.
- [265] N. PAPAMARKOS and B. GATOS, A new approach for multilevel threshold selection, CVGIP-IP(56), No. 5, September 1994, pp. 357-370.
- [266] P. PEREZ et F. HEITZ, Une approche multi-échelle l'analyse d'images par champs markoviens. TS, 9(6) pp. 459-472, 1992.
- [267] W. PIECZYNSKI, Statistical image segmentation, Machine Graphics and Vision, Proc. of GKPO'92, Vol. 1, no. 1/2, pp. 261-268, Mai 1992.
- [268] W. PIECZYNSKI, Champs de Markov cachés et estimation conditionnelle itérative. TS, 11(2) pp. 141-153, 1994.
- [269] M. PETTROU and J. KITTLER, Optimal edge detectors for ramp edges, T-PAMI-13(5) pp. 483-491,1991.
- [270] T. PAVLIDIS and Y.T. LIOW, Integration region growing and edge detection, T-PAMI-12(3) pp. 225-233,1990.

- [271] P. PROVENT, J. LEMOINE et E. PETIT, Segmentation optimale d'images de textures par transformations adaptatives multiples, TS 11(1) pp. 43-56,1994.
- [272] J.M.S. PREWITT and M.L. MENDELSON, The analysis cell images, in Ann. New York Acad. Sci., Vol 128, pp. 1035-1053, New York Acad. Sci., New York, 1966.
- [273] P. PERONA and J. MALIK, Scale-Space and edge detection using anisotropic diffusion, T-PAMI-12(7) pp. 629-639,1989.
- [274] P. PERONA and J. MALIK, Detecting and localizing edges composed of steps, peaks and roofs, In ICCV, Osaka Dec. 1990, pp. 52-57.
- [275] J. PARKKINEN and E. OJA, Cooccurrence matrices and subspace methods in texture analysis. 8ème ICPR, pp. 405-408, Paris, France, Octobre 1986.
- [276] S.K. PAL and N.R. PAL, Segmentation using contrast and homogeneity measures, PRL 5 pp. 293-304, 1987
- [277] S.K. PAL and N.R. PAL, Segmentation Based on Measures of Contrast, Homogeneity, and Region Size. IEEE Trans. Systems, Man, and Cybernetics, SMC 17(5) pp. 857-868, Sept. 1987.
- [278] S.K. PAL and N.R. PAL, Segmentation using contrast and homogeneity measures, PRL 5 pp. 293-304,1987.
- [279] N.R. PAL and S.K. PAL, Entropic thresholding, SP 16(2) pp. 97-108,1989.
- [280] N.R. PAL and S.K. PAL, A review of image segmentation techniques. PR, 26(9) pp. 1277-1294, 1993.
- [281] S. PELEG and A. ROSENFELD, Determining compatibility coefficients for curve enhancement relaxation processes, IEEE Trans. Syst. Man. Cybern., Vol. SMC-8, n1, 1978.
- [282] S.K. PAL and A. ROSENFELD, Image enhancement and thresholding by optimization of fuzzy compactness. PRL, (7) pp. 77-86, 1988.
- [283] W.P. Pratt, Digital Image Processing, Wiley Interscience Publication, NY 1978.
- [284] J.M.S. PREWITT, Object enhancement and extraction, In Picture Processing and Psychopicturics, B.S.Lipkin and A.Rosenfeld Eds. Academic Press NY. 1970.
- [285] J. PARKKINEN, K. SELKAINAHO and E. OJA, Detecting texture periodicity from the cooccurrence matrix, PRL 11(1) pp. 43-50,1990.
- [286] T. PUN, A new method for gray-level picture thresholding using the entropy of the histogram, Signal Process. 2, pp. 223-237, 1980.
- [287] T. PUN, Entropic thresholding: A new approach CVGIP. 16, pp. 210-239, 1981.
- [288] P. PARENT and S.W. ZUCKER, Trace inference, curvature consistency, and curve detection. T-PAMI, 1(8), Aug. 1989.
- [289] T.R. REED and J.M. HAS DU BUF, A review of recent texture segmentation and feature extraction techniques. CVGIP: IU, 57 pp. 359-372, Mai 1993.
- [290] C. ROUQUET and P. BONTON. Region-based segmentation of textured images. 8th IAPR, pp. 11-16, San Remo, Italie, Septembre 1995.
- [291] A. RANGARAJA and R. CHELLAPPA, Generalise graduated non convexity algorithm for MAP image estimation, Proc. of ICPR June 1990, pp. 127-133.
- [292] A. ROSENFELD, and L.S. DAVIS, Iterative histogram modification, T-SMC-8 pp. 300-302,1978.
- [293] A. ROSENFELD and P. De La TORRE, Histogram concavity analysis as an aid in threshold selection, T-SMC-13 pp. 231-235,1983.
- [294] A. ROSENFELD and A. KAK, Digital Image Processing, volume 2. Academic Press, New York, 1982.
- [295] M. ROUX and D.M. MCKEOWN, Feature matching for building extraction from multiple views. IEEE CVPR, pp. 46-53, Jun. 1994.

- [296] L.G. ROBERTS, Machine perception of three-dimensional solids, chaoter optical and electro-optical information processing. J.P.Tippet and al. Ed. MIT Press, pp. 159-197, Cambridge 1965.
- [297] C. ROBA et J.G. POSTAIRE, Une transformation non linéaire de l'histogramme des niveaux de gris pour la sélection automatique des seuils de segmentation d'images, 13ième colloque GRETSI, Juan-les-Pins, pp. 1093-1095, Sept. 1991.
- [298] R.S. PELEG and A. ROSENFELD, Shape segmentation using relaxation, T-PAMI, Vol 3, pp. 368-375, July 1981.
- [299] S.S. REDDI, S.F. RUDIN, and H.R. KESHAVAN, An optimal multiple threshold scheme for image segmentation, T-SMC, SCM-14;661-665, 1984.
- [300] A. ROSENFELD and R.C. SMITH, Thresholding using relaxation, T-PAMI-3 pp. 598-606, 1981.
- [301] A. ROSENFELD and P. De La TORRE, Histogram concavity analysis as an aid in threshold selection, T-SMC-13 pp. 231-235, 1983.
- [302] T. REED and H. WECHSLER, Tracking for non stationarities for texture fields, SP 14(1) pp. 95-102,1988.
- [303] T. REED, H. WECHSLER and M. WERMAN, Texture segmentation using a diffusion region growing technique, PR 23(9) pp. 953-960,1990.
- [304] H. SAMET, Connected composants labelling using quadtree, TR-756, Univ. of Maryland, 1979.
- [305] F. SANDAKLY, Contribution à la mise en oeuvre d'une architecture à base de connaissance pour l'interprétation de scènes 2D et 3D. Thèse de l'université de Nice-Sophia Antipolis, 1995.
- [306] S. SARKAR and K.I. BOYER, On optimal impulse response edge detection filter, T-PAMI-13(11) pp. 1154-1171,1991.
- [307] B. SERRA and M. BERTHOD, Subpixel contour matching using continuous dynamic programming. In CVPR, pp. 202-207, Seattle, 1994.
- [308] J. SHEN and S. CASTAN, An optimal linear operator for step edge detection. CVGIP, 54(2) pp. 112-133, March 1992.
- [309] P. SAINT-MARC, J.S. CHEN, and G. MEDIONI, Adaptative smoothing: a general tool for early vision. T-PAMI-13(6) pp. 514-528,1991.
- [310] J. SERRA, Image Analysis and Mathematical Morphology, Academic Press London 1982.
- [311] M. SPANN and A.E. GRACE, Adaptative segmentation of noisy and textured images. PR, 27(12) pp. 1717-1733, 1994.
- [312] G. SHAFER, Implementing dempster's rule for hierarchical evidence. In AI, volume 33, pp. 271-298. 1987.
- [313] G. SHAFER, A Mathematical Theory of Evidence. Princeton University Press, 1979.
- [314] R. SHAPLEY and D.M. LAM, Contrast Sensitivity. Cambridge, MA, The MIT Press, 1993.
- [315] S. SANTINI and R. JAIN, Similarity Matching, Technical Repport, Visual Computing Laboratory, University of California, San Diego, 1996.
- [316] I. SOBEL, An isotropic image gradient operator. In H. Freeman, editor, Machine Vision for Three-Dimensional Scenes, pp376-379. Academic Press, 1990.
- [317] R. SOUTHWELL, Relaxation methods in theoretical physics, Oxford Univ. Press (Clarendon), London, 1946.
- [318] B.J. SCHARCHTER, A. ROSENFELD, and L.S.DAVIS, Random mosaic models for textures, T-SMC-8(9) pp. 694-702,1978.
- [319] P.K. SAHOO, S. SOLTANI, A.K.C. WONG and Y.C. CHEN A survey of thresholding techniques, CVGIP. Vol 41, pp. 233-260, 1988.
- [320] S. STRYHANYN-CHARDON, Contribution au problèmes de la restauration myope d'images numériques: analyse et synthèse. Thèse de doctorat en cours, Université de Rennes I.
- [321] J. SERRA and G. VERCHARY, Mathematical morphology applied to fibre composite materials, Film Sci. Tech. 6 pp. 141-158,1973.

- [322] C. SUN and W.G. WEE, Neighbouring gray level dependence matrix for texture classification, CVGIP 23 pp. 341-352,1983.
- [323] M. SPANN and R. WILSON, A quadtree approach to image segmentation which combines statistical and spatial information. PR, (3/4) pp. 257-269, 1985.
- [324] A.N. TIKHONOV et V.Y. ARSENIN, Méthodes de résolution de problèmes mal posés, MIR Moscou 1974.
- [325] M.J. TARR and M.J. BLACK, A Computational and Evolutionary Perspective on the Role of Representation in Vision. CVGIP: IU, 60(1) pp. 65-73, July 1994.
- [326] C.H. TEH and R.T. CHIN, On digital approximation of moment invariants, CVGIP 33, pp318-326, 1986.
- [327] C.H. TEH and R.T. CHIN, On image analysis by the methods of Moments, T-PAMI, Vol 10, n4, pp . 496-512, July 1988.
- [328] A.M. TREISMAN and G. GELADE, A feature-integration theory of attention. Cognitive psychology, 12 pp. 97-136, 1980.
- [329] H.L. TAN, S.B. GELFAN, and E.J. DELP, A cost minimization approach to edge detection using simulated annealing. T-PAMI, 14 pp. 3-18, jan 1991.
- [330] M. TUCERYAN and K. JAIN, Texture segmentation using Voronoi polygons. T-PAMI, (2) pp. 211- 216, 1990.
- [331] M. TUCERYAN, A.K. JAIN and Y. LEE, Extracting perceptual structures in line patterns. In Proc. of 5th Scandinavian Conference on Image Analysis, Stockholm (Sweden), pp. 531-537, 1987.
- [332] V. TORRE and T. POGGIO, On edge detection, T-PAMI-8 pp. 679-698,1986.
- [333] W. TSAI, Moment-preserving thresholding: a new approach, CVGIP. 29, pp. 377-393, 1985.
- [334] J.K. TSOTSOS, There is no one way to look at vision. CVGIP: IU, 60(1) pp. 95-97, July 1994.
- [335] M.R. TURNER, Texture discrimination by Gabor functions, Biological. Cybern. 55 pp. 71-82,1986.
- [336] S. TABBONE et D. ZIOU, Description multi-échelle d'une image fondée sur le comportement de contour. Actes de Cong. RFIA Lyon 1991, pp. 567-576.
- [337] M. UNSER and M. EDEN, Non-linear operators for improving texture segmentation based on features extracted by spatial filtering, T-PAMI-20(4) pp. 804-815,1990.
- [338] S. ULLMAN, Visual routines. MIT Press, Cambridge 1983.
- [339] M. UNSER, Local linear transforms for texture measurements, SP 11 pp. 61-79,1986.
- [340] L. VINET et al. Cooperative segmentation and stereo matching, Image Topical Meeting on IU and Machine Vision, Massachusetts, June 1989.
- [341] L. VINET, Segmentation et mise en correspondance de régions de paires d'images stéréoscopiques. Thèse, Université de Dauphine, Paris IX, Juillet 1991.
- [342] S. VENKATESH and L.J. KITCHEN, Edge evaluation using necessary components. CVGIP, 54(1) pp. 23-30, January 1992.
- [343] P.J. VAN LAARHOVEN and E.H. AARTS, Simulated annealing: Theory and application, Riddle Pub. Dordrecht 1987.
- [344] F.M. VILNROTTER, R. NEVATIA and K.E. PRICE, structural analysis of natural textures, T-PAMI-8(1) 1986.
- [345] P. VIRGILI, G. VOTTERO, G. VERNAZZA, and S. DELLEPIANE. A segmentation method for fuzzy boundary detection. In Proc. of the 6th International Conference on Image Analysis and Processing, pp. 60-64, Como, Italy, september 1991.
- [346] E. WALTZ and J. LLINAS. Multisensor data fusion. Artech House, Boston, 1990.
- [347] S. WATANABE, and CYBEST group, An automated apparatus for cancer processing CYBEST, CVGIP. 3, pp. 350-358, 1974.

- [348] A.B. WATSON, Digital images and human vision. Cambridge MA, MIT Press, 1993.
- [349] C.H. WU and P.C. DOERSSCHUK. Tree Approximations to Markov Random Fields. T-PAMI 17(4) pp. 391-402, Avril 1995.
- [350] J.S. WESKA, A survey of threshold selection techniques, CVGIP. 7, pp. 259-265, 1978.
- [351] S. WANG and R.M. HARALICK, Automatic multithreshold selection, Comput; Vision Graphics Image Process. 25, pp. 46-67, 1984.
- [352] L. WANG and D.C. HE, Texture classification using texture spectrum, PR 23(8) pp. 905-909, 1990.
- [353] A.Y. WU, T-H. HONG and A. ROSENFELD, Threshold selection using quaternaries, T-PAMI-4, pp. 90-94, January 1982.
- [354] R. WILSON. H. KNUTSSON and G.H. GRANLUND, The operational definition of the position of line and edge Proc. ICPR-82, Munich 1982.
- [355] B. WROBEL et O. MONGA, Segmentation d'images naturelles: coopération entre un détecteur-contour et un détecteur-région, Actes du 11ème colloque GRETSI, Nice, pp. 539-542, 1987.
- [356] B. WROBEL et O. MONGA, Segmentation d'images naturelles: Coopération entre un détecteur de contour et un détecteur de région, Proc. of Cong. COGNITIVA MARI, pp. 181-188, Paris Mai 1987.
- [357] J.S. WESKA, R.N. NAGEL, and A. ROSENFELD, A threshold selection technique, IEEE Trans. Comput. C-23, pp. 1322-1326, 1974.
- [358] R.N. WOLFE, A dynamic thresholding scheme for quantization of scanned image, in Proceedings, Automatic PR, pp. 143-162, 1969.
- [359] G. WOLBERG and T. PAVLIDIS, Restoration of binary images using stochastic relaxation with annealing. PRL, 3 pp. 375-388, Décembre 1985.
- [360] J.S. WESKA and A. ROSENFELD, Threshold evaluation techniques, T-SMC-8, pp. 622-629, 1978.
- [361] J.S. WESKA and A. ROSENFELD, Histogram modification for threshold selection, T-SMC-9, pp. 38-51, 1978.
- [362] J.M. WHITE and G.D. ROHRER, Image thresholding for optical character recognition and other applications requiring character image extraction, IBM J. Res. Development, 27, pp. 400-411, 1986.
- [363] X. WU, Adaptive split-and-merge segmentation based on piecewise least-square approximation, PAMI(15), No. 8, pp. 808-815, August 1993.
- [364] T. WANG, X. ZHUANG and X. XING, Robust and adaptive segmentation of noisy images using Gibbs random fields models. T-PAMI, 6(2) pp. 753-775, 1992.
- [365] Y. XIAOHAN and al. Image segmentation combining region growing and edge detection, Proc. of the ICPR, The Hague Nd. Sep. 1992.
- [366] X.L. XIE and G. BENI, A validity measure for fuzzy clustering. T-PAMI, (8) pp. 841- 847, 1991.
- [367] W.A. YASSNOF, J.K. MUI, and J.W. Bacus, Error measures for scene segmentation, PR, vol. 9, 217-231, 1977.
- [368] Y.J. ZHANG and J.J. GERBRANDS, Segmentation evaluation using ultimate measurement accuracy. Image Processing Algorithms and Techniques III, SPIE 1657 pp. 449-460, 1992.
- [369] Y.J. ZHANG and J.J. GERBRANDS, Comparison of thresholding techniques using syntactic images and ultimate measurements accuracy, Proc. of the ICPR The Hague Nd Sep. 1992.
- [370] Y.J. ZHANG, "A survey on evaluation methods for image segmentation", PR, Vol.29, No 8, pp. 1335-1346, 1996.
- [371] S. ZUCKER, R. HUMMEL, and A. ROSENFELD, An application of relaxation labelling to line and curve enhancement, IEEE Trans. Comput. C-26, pp. 394-403, 1977.
- [372] D. ZIOU, Line detection using an optimal IIR filter, PR 24(6) pp. 465-478, 1991.

- 
- [373] J. ZHANG and H.H. NAGEL, Texture based segmentation of road images. the Intelligent Vehicules, pp. 24-26, Paris, France, Octobre 1994.
- [374] J. ZERUBIA et F. PLOYETTE, Détection de contour et lissage d'image par deux algorithmes déterministes de relaxation, TS 8(3) pp. 165-179, 1990.
- [375] C.T. ZAHN and R.Z. ROSKIES, Fourier descriptions of plane closed curves, IEEE Trans. on Computers Anal., vol C-21, pp . 269-281, 1972.
- [376] S. ZUCKER, Survey: region growing, childhood and adolescence. CVGIP, 5 pp. 382-399, 1976.
- [377] Y.T. ZHOU, V.VENKATESWAR and R.CHELLAPA, Edge detection and linear feature extraction using a 2D random field model, T-PAMI-11(1) pp. 84-94, 1989.
- [378] S.C. ZHU and A. YUILLE, Region competition: unifying snakes, region growing and bayes/mrf for multiband image segmentation, IEEE PAMI-18(9), Sept. 1996.

## Liste des tableaux

3.1	Comportement de la procédure de regroupement pour l'image Aquitaine en fonction de la taille de la fenêtre de la moyenne, $m$ . . . . .	130
3.2	Comportement de la procédure de regroupement pour l'image Savoie lissée en fonction de la taille de la fenêtre de la moyenne, $m$ . . . . .	130
3.3	Comportement de la procédure de regroupement pour l'image Paimpont en fonction de la taille de la fenêtre de la moyenne, $m$ . . . . .	130
3.4	Application des différentes méthodes de multi-seuillage sur l'image Savoie. $W$ : taille de fenêtre, $NC$ : nombre de classes, $NV$ : nombre de voisins, $NS$ : nombre de seuils. . . . .	136
3.5	Dissimilarités calculées entre les différents résultats de multi-seuillage $RS_i$ de l'image Savoie et la segmentation de référence $Rf$ . $W$ : taille de fenêtre, $NC$ : nombre de classes, $NV$ : nombre de voisins, $NS$ : nombre de seuils. . . . .	136
3.6	Application des différentes méthodes de multi-seuillage sur l'image Paimpont. $W$ : taille de fenêtre, $NC$ : nombre de classes, $NV$ : nombre de voisins, $NS$ : nombre de seuils. . . . .	140
3.7	Dissimilarités calculées entre les différents résultats de multi-seuillage $RS_i$ de l'image Paimpont et la carte vérité de terrain $V_1$ . $NC$ : nombre de classes. . . . .	140
3.8	Mesures selon l'uniformité des niveaux de gris par région. Image Savoie bruitée.	142
3.9	Mesures selon l'uniformité des niveaux de gris par niveau de gris. Image Savoie bruitée. . . . .	142
3.10	Mesures selon l'uniformité des niveaux de gris par région. Image Paimpont. . . . .	143
3.11	Mesures selon l'uniformité des niveaux de gris par niveau de gris. Image Paimpont. . . . .	143
3.12	Statistiques de 20 classes. . . . .	158
3.13	Statistiques de 12 classes. . . . .	158
3.14	Taux d'identification en (%) entre les classes de la vérité de terrain (première colonne) et les classes issues de la modélisation par les co-occurrences du résultat obtenu par la méthode CRT (première ligne). . . . .	162
3.15	Taux d'identification en (%) entre les classes issues de la modélisation par les co-occurrences du résultat obtenu par la méthode CRT (première ligne) et les classes de la vérité de terrain (première colonne). . . . .	162

3.16	Taux d'identification en (%) entre les classes de la vérité de terrain (première colonne) et les classes issues de la modélisation par les longueurs de plages du résultat obtenu par la méthode CRT (première ligne). . . . .	163
3.17	Taux d'identification en (%) entre les classes issues de la modélisation par les longueurs de plages du résultat obtenu par la méthode CRT (première ligne) et les classes de la vérité de terrain (première colonne). . . . .	163
3.18	Taux d'identification en (%) entre les classes de la vérité de terrain (première colonne) et les classes issues de la modélisation par les deux premiers moments du résultat obtenu par la méthode CRT (première ligne). . . . .	164
3.19	Taux d'identification en (%) entre les classes issues de la modélisation par deux premiers moments du résultat obtenu par la méthode CRT (première ligne) et les classes de la vérité de terrain (première colonne). . . . .	164
3.20	Dissimilarités calculées entre les résultats obtenus par le système de segmentation (cas contour-région-texture CRT et cas contour-région CR) et la carte vérité de terrain <i>VT</i> . . . . .	165
3.21	Taux d'identification en (%) - algorithme MAP-YOUNES. . . . .	167
3.22	Taux d'identification en (%) - algorithme SEM adaptatif MAP. . . . .	167

# Table des figures

2.1	Classification des méthodes par seuillage . . . . .	18
2.2	Exemple d'une décomposition spatiale et "quadtree" correspondant. . . . .	38
2.3	Echantillons de textures pris de l'album de Brodatz. . . . .	52
2.4	Classification des méthodes de segmentation de textures . . . . .	53
2.5	Vecteurs de déplacement. . . . .	54
2.6	Causalité dans les modèles 2D . . . . .	61
2.7	Système de voisinage et cliques associées . . . . .	64
3.1	Contour de type "marche" bruité. SNR=1 à gauche et SNR=0.25 à droite [108].	88
3.2	Filtrage de la "marche" bruitée avec un SNR=1. $\alpha=.25$ ; $\alpha=.5$ ; $\alpha=.75$ ; $\alpha=1$ .	88
3.3	Filtrage de la "marche" bruitée avec un SNR=0.25. $\alpha=.25$ ; $\alpha=.5$ ; $\alpha=.75$ ; $\alpha=1$ .	88
3.4	(a) Image SPOT de la région Aquitaine après réhaussement. Détection des maxima locaux par filtre de Deriche. (b) $\alpha=0.5$ ; (c) $\alpha=1$ . (d) $\alpha=2$ . Les points en noir correspondent aux maxima des gradients non nuls. . . . .	89
3.5	Extraction d'un noyau homogène. . . . .	94
3.6	Extraction de régions par agrégation récursive de points appliquée à l'image Aquitaine. (a) $S_k=6$ . (b) $S_k=7$ . (c) $S_k=8$ . (d) $S_k=9$ . . . . .	97
3.7	Synoptique du système de segmentation: cas de la coopération contours-régions uniformes. . . . .	99
3.8	Croissance de région autour de noyaux homogènes guidée par les contours. .	101
3.9	Extraction des maxima locaux. . . . .	102
3.10	Correction des irrégularités locales. . . . .	103
3.11	Groupement perceptuel des contours. . . . .	104
3.12	Sensibilité de la distance de Hausdorff au bruit. . . . .	108
3.13	Espace de recherche du point le plus proche. . . . .	111
3.14	Segmentations en contours et en régions intermédiaires, image Quart-Aquitaine.	112
3.15	Suite, segmentations intermédiaires et finales en contours et en régions, image Quart-Aquitaine. . . . .	113
3.16	Evolution de la dissimilarité entre la carte des contours et la carte des frontières de régions en fonction des itérations. Image Quart-Aquitaine. . . . .	114
3.17	Segmentation coopérative contours-régions. (a) Image Aquitaine. (b) Résultat final de la segmentation région. . . . .	115

3.18	Segmentation coopérative contours-régions. (a) Carte de contours de départ. (b) Frontières finales de la segmentation région. . . . .	115
3.19	Evolution de la dissimilarité entre la carte des contours et la carte des frontières de régions en fonction des itérations. Image Aquitaine. . . . .	116
3.20	Segmentation coopérative contours-régions. Image IRM. . . . .	116
3.21	Segmentation coopérative contours-régions. Image Car. . . . .	117
3.22	Segmentation coopérative contours-régions. Image Aérienne. . . . .	117
3.23	Synoptique du système de segmentation: cas de la coopération Contours-Régions-Textures. . . . .	118
3.24	Images de bandes verticales de niveaux de gris ayant servi à la construction de fonction de la sensibilité de l'oeil au contraste. a) Bandes contrastées avec un $C = 4$ . b) $C = 8$ . c) $C = 20$ . . . . .	121
3.25	Fonction de perception visuelle . . . . .	122
3.26	a) Image Savoie bruitée. b) Son histogramme. . . . .	124
3.27	Histogramme global des maxima locaux $H_{glob}$ de l'image Savoie bruitée. . .	124
3.28	a) Image SPOT de la forêt de Paimpont. b) Son histogramme. . . . .	125
3.29	Histogramme global des maxima locaux $H_{glob}$ de l'image de la forêt de Paimpont.	125
3.30	Masses ponctuelles après groupement des masses de $H_{glob}$ . (a) de l'image Savoie bruitée. (b) l'image de la forêt de Paimpont. . . . .	126
3.31	Image Paimpont. Seuils de segmentation: 44 77 100 123 148 176 195 215 . .	127
3.32	Image IRM. Seuils de segmentation: 17 45 72 94 107 123 153 184 207 222 . .	128
3.33	Image Car. Seuils de segmentation: 15 35 55 80 111 135 155 176 194 . . . . .	129
3.34	a) Image Savoie bruitée de la banque d'image du GDR TdSI. b) Image Référence Savoie $Rf$ . . . . .	134
3.35	Résultats de segmentation de l'image Savoie bruitée, obtenus par les différentes méthodes de multi seuillage. a) Résultat de segmentation obtenu par la méthode développée. b) Méthode THi. c) Méthode AD. d) Méthode des Kmeans. e) Méthode MC1 avec $NV = 20$ . f) Méthode MOM avec $NC = 3$ . .	134
3.36	Résultats de segmentation de l'image Savoie bruitée, obtenus par les différentes méthodes de multi seuillage. a) Méthode MT. b) Méthode d'entropie avec $NV = 2$ . c) Méthode d'entropie avec $NV = 6$ . d) Méthode MC2 avec $NV = 1$ . e) Méthode MC2 avec $NV = 8$ . f) Méthode LdP $NV = 23$ . g) Méthode CM. . . . .	135
3.37	(a) Image de la forêt de Paimpont masquée. (b) Contours vérité de terrain de la forêt de Paimpont. Image $VT$ . . . . .	138
3.38	Résultats de segmentation de l'image Paimpont obtenus par les différentes méthodes de multi seuillage. a) Résultat de segmentation obtenus par la méthode développée. b) Méthode THi. c) Méthode AD. d) Méthode des Kmeans. e) Méthode MC1 avec $NV = 1$ . f) Méthode MC1 avec $NV = 2$ . . . . .	138

3.39	Résultats de segmentation de l'image Paimpont obtenus par les différentes méthodes de multi seuillage. a) Méthode MT. b) Méthode d'entropie avec $NV = 5$ . c) Méthode d'entropie avec $NV = 6$ . d) Méthode MC2 avec $NV = 1$ . e) Méthode MC2 avec $NV = 2$ . f) Méthode LdP. g) Méthode CM. . . . .	139
3.40	Voisinage utilisé pour la segmentation en régions. (Avec chevauchement spatial).	147
3.41	Histogramme identique pour des textures différentes. . . . .	148
3.42	(a) Image originale composée des textures D24, D29, D38 et D68. (b) Résultat du multi-seuillage de l'image de textures. (c) Résultat de la segmentation par fusion de régions après un filtrage médian, avec une fenêtre de taille 32x32, et un seuil de fusion égal à 35% de la surface de la fenêtre. (d) Avec une fenêtre de taille 16x16, et un seuil de fusion égal à 15% de la surface de la fenêtre. . .	152
3.43	(a) Méthode basée sur les histogrammes d'ordre 2. Résultat de la segmentation par fusion de régions après un filtrage médian, avec une fenêtre de taille 16x16, et un seuil de fusion égal à 40% de la surface de la fenêtre. (b) Résultat de la segmentation par fusion de régions après un affinement par modélisation AR.	153
3.44	Interpolation des points. . . . .	155
3.45	Segmentation obtenus après intégration de la texture dans la coopération contours-régions: l'image Aérienne. a) Son histogramme. b) Son histogramme global $H_{glob}$ . c) Son histogramme global après regroupement des masses ponctuelles. d) Résultat du multi-seuillage (seuils: 98, 119, 135, 153, 176, 201). e) Résultat de la fusion des régions multi-seuillées. f) Segmentation après intégration de la texture. . . . .	156
3.46	Image de la forêt de Paimpont masquée. . . . .	159
3.47	Carte vérité de terrain de la forêt de Paimpont VT. En contours (à gauche) et en régions (à droite). . . . .	159
3.48	Contours de la carte vérité de terrain de la forêt de Paimpont superposés à l'image originale. . . . .	160
3.49	Résultats de segmentation de l'image Paimpont obtenus par la méthode coopérative, (a) contours-régions (CR); (b) contours-régions-textures (CRT); développée superposée à l'image originale. . . . .	160
3.50	Résultat de segmentation de l'image Paimpont obtenu par la méthode CRT développée après classification en 12 classes en utilisant (a) les attributs de co-occurrence; (b) les attributs de longueurs de plages. . . . .	161
3.51	Résultat de segmentation de l'image Paimpont obtenu par la méthode CRT après classification en 12 classes en utilisant les 2 premiers moments. . . . .	161
3.52	Contours du résultat de segmentation de l'image Paimpont obtenu par: (a) la méthode coopérative contours-régions (méthode CR); (b) la méthode coopérative contours-régions-textures (méthode CRT). . . . .	165
3.53	Résultat de segmentation de deux méthodes markoviennes. (a) algorithme MAP-YOUNES. (b) algorithme SEM adaptatif MAP. . . . .	166
A.1	Les opérateurs de lissage et de dérivation. C1: exponentiel; C2: gaussienne . .	178