

N° d'ordre:

# THÈSE

présentée à

**L'UNIVERSITÉ DE NICE SOPHIA ANTIPOLIS**

pour obtenir le titre de

**DOCTEUR EN SCIENCES**

Spécialité

**SCIENCES DE L'INGÉNIEUR**

par

**Anne LORETTE**

Sujet de la thèse:

**ANALYSE DE TEXTURE PAR MÉTHODES  
MARKOVIENNES ET PAR MORPHOLOGIE  
MATHÉMATIQUE : APPLICATION À L'ANALYSE DES  
ZONES URBAINES SUR DES IMAGES SATELLITALES**

Soutenue le 28 Septembre 1999 devant le jury composé de :

Mr.	<b>Albert</b>	<b>Bijaoui</b>	Président
Mr.	<b>Donald</b>	<b>Geman</b>	Rapporteurs
Mr.	<b>Hervé</b>	<b>Le Men</b>	
Mr.	<b>Henri</b>	<b>Maitre</b>	
Mr.	<b>Thierry</b>	<b>Phulpin</b>	Examineurs
Mr.	<b>Xavier</b>	<b>Descombes</b>	
Mme	<b>Josiane</b>	<b>Zerubia</b>	



# Remerciements

Je remercie Mr. Bijaoui qui a accepté de présider ce jury.

J'adresse des remerciements tout particulier à Mr. Geman, Mr. Maître, et Mr. Le Men, pour le temps qu'ils ont consacré à la relecture du manuscrit ainsi que pour leurs remarques constructives.

Je suis reconnaissante à Mme Zerubia de m'avoir accueillie au sein de son laboratoire et d'avoir dirigé mes travaux, ainsi qu'à Xavier Descombes pour l'aide scientifique qu'il m'a apporté.

Je tiens également à remercier le personnel de la documentation, du Semir, et le personnel administratif dans son ensemble.

Enfin un très grand merci à mes amis, mes parents, mon frère et ma soeur.



# Table des matières

<b>Table des matières</b>	<b>i</b>
<b>Figures</b>	<b>i</b>
<b>Tableaux</b>	<b>v</b>
<b>Introduction</b>	<b>1</b>
<b>1 Contexte - Objectifs - Méthodes</b>	<b>5</b>
1.1 Contexte Général . . . . .	6
1.2 Contexte Particulier . . . . .	9
1.2.1 Données . . . . .	10
1.2.2 Notations et Définitions . . . . .	12
1.2.2.1 Inférence bayésienne . . . . .	13
1.2.2.2 Distribution a priori et a posteriori . . . . .	13
1.2.2.3 Risque et estimateur bayésien . . . . .	13
1.2.2.4 Champs de Markov . . . . .	14
1.2.2.5 Champs de Gibbs . . . . .	15
1.2.3 Analyse de texture : état de l'art . . . . .	16
1.2.3.1 Méthodes statistiques et spectrales . . . . .	17
1.2.3.2 Méthodes géométrique . . . . .	17
1.2.3.3 Méthodes de filtrage . . . . .	18

1.2.3.4	Méthodes fondées sur un modèle . . . . .	20
1.2.3.5	Méthodes mixtes . . . . .	23
1.2.3.6	Conclusion . . . . .	23
<b>2</b>	<b>Masque Urbain: Analyse de texture</b>	<b>25</b>
2.1	Introduction . . . . .	26
2.2	Discrimination bâti/non bâti: état de l'art et applications . . . . .	26
2.3	Modèles isotropes . . . . .	29
2.3.1	Modèle markovien gaussien . . . . .	29
2.3.1.1	Définition . . . . .	29
2.3.1.2	Estimation des paramètres : "queues de comètes" . . . . .	30
2.3.2	Matrices aura . . . . .	31
2.3.3	Pourquoi prendre en compte l'anisotropie ? . . . . .	37
2.4	Modèles fondés sur des chaînes . . . . .	39
2.4.1	Introduction . . . . .	39
2.4.2	Les modèles . . . . .	40
2.4.3	Normalisation des paramètres . . . . .	42
2.4.4	Résultats . . . . .	50
2.4.5	Comparaison avec deux méthodes classiques . . . . .	53
2.4.5.1	Filtres de Gabor . . . . .	53
2.4.5.2	Matrices de cooccurrence . . . . .	64
2.4.6	Conclusion . . . . .	69
<b>3</b>	<b>Masque Urbain: Segmentation</b>	<b>71</b>
3.1	Classification floue . . . . .	72
3.1.1	Ensembles flous . . . . .	72
3.1.2	Critère classique . . . . .	72
3.1.3	Critère entropique . . . . .	74
3.1.4	Résultats expérimentaux . . . . .	79

3.2	Régularisation par modélisation markovienne . . . . .	81
3.2.1	Modèle de segmentation . . . . .	87
3.2.1.1	Critère du MAP . . . . .	87
3.2.1.2	Vraisemblance . . . . .	87
3.2.1.3	A priori . . . . .	89
3.2.2	Algorithme ICM ( <i>Iterated conditional modes</i> ) . . . . .	89
3.2.3	Image de marqueurs . . . . .	90
3.3	Résultats . . . . .	90
3.4	Conclusion . . . . .	94
<b>4</b>	<b>Analyse Intra-Urbaine</b>	<b>111</b>
4.1	Introduction . . . . .	112
4.2	Analyse intra-urbaine: quelques applications . . . . .	112
4.3	Généralités et définitions . . . . .	114
4.3.1	Morphologie mathématique : quelques généralités . . . . .	114
4.3.2	Les transformations de base : définitions et propriétés . . . . .	115
4.3.2.1	Définitions . . . . .	115
4.3.2.2	Propriétés algébriques . . . . .	118
4.3.2.3	Propriétés géométriques . . . . .	119
4.3.3	Ouverture par Reconstruction . . . . .	119
4.3.3.1	Introduction . . . . .	119
4.3.3.2	Transformations géodésiques . . . . .	120
4.3.3.3	Ouverture par reconstruction . . . . .	121
4.3.3.4	Ouverture par reconstruction versus ouverture morphologique : exemples dans les cas binaire et en niveaux de gris . . . . .	123
4.4	Granulométrie et distribution de taille . . . . .	124
4.4.1	Introduction . . . . .	124
4.4.2	Granulométrie . . . . .	127

4.4.3	Distribution de taille . . . . .	128
4.4.4	Application à l'analyse intra-urbaine . . . . .	129
4.4.4.1	Granulométrie . . . . .	129
4.4.4.2	Quantification . . . . .	131
4.5	Modèle markovien de segmentation . . . . .	131
4.5.1	Introduction . . . . .	131
4.5.2	Le modèle . . . . .	132
4.5.3	Algorithme du recuit simulé . . . . .	133
4.5.4	Segmentation : résultats expérimentaux . . . . .	134
4.6	Conclusion . . . . .	137
	<b>Conclusion</b>	<b>143</b>
	<b>Annexe</b>	<b>147</b>
	Renormalisation par décimation . . . . .	147
	<b>Bibliographie</b>	<b>149</b>

# Figures

<b>Chapitre 1.</b>	<b>5</b>
1.1 Exemple de simulations SPOT5 : Toreilles. . . . .	12
<b>Chapitre 2.</b>	<b>25</b>
2.1 Queue de comète. . . . .	32
2.2 Exemple d'un ensemble <code>aura</code> (b) sur un réseau de sites binaires (a) pour un système de voisinage 4 connexe. . . . .	33
2.3 Image initiale : Toreilles. . . . .	37
2.4 Résultats comparatifs sur l'image de Toreilles : (a) entropie, (b) énergie, (c) variance conditionnelle locale, (d) variance. . . . .	38
2.5 Zooms sur l'image de Toreilles : structures anisotropes (serres, cultures linéaires). . . . .	39
2.6 Représentation des 8 directions prises en compte. . . . .	42
2.7 Représentation du réseau complet et des réseaux décimés. . . . .	45
2.8 Fonctions de correction $h(\lambda)$ et $g(\lambda)$ . . . . .	49
2.9 Texture isotrope (512×512) : surface d'une feuille de papier. . . . .	50
2.10 Fonctions de répartition du paramètre de texture estimé dans chacune des 8 directions. . . . .	51
2.11 Images des estimées (fenêtre de taille 17×17) des paramètres de texture (vidéo inversée) sur l'image de Toreilles. . . . .	54
2.12 Images des estimées (fenêtre de taille 17×17) des paramètres de texture (vidéo inversée) sur l'image de Toreilles. . . . .	55
2.13 Rose des directions sur des images de champs extraites d'une simulation SPOT5 (en haut) et d'une image SPOT3 (en bas). . . . .	56

2.14	Rose des directions sur des images ayant une orientation privilégiée extraites d'une simulation SPOT5. . . . .	57
2.15	Rose des directions sur des images de zones urbaines extraites d'une images SPOT3 (image du haut) et de simulations SPOT5 (les deux images du bas). . . . .	58
2.16	Rose des directions sur des images de zones urbaines extraites de simulations SPOT5. . . . .	59
2.17	Rose des directions sur des images extraites d'images ERS1 (haut : champs, les 2 images du bas : ville). . . . .	60
2.18	(a): Image SPOT3 de la ville de Cayenne, (b): Estimée de la variance conditionnelle locale du modèle isotrope 2D, (c): Estimée du paramètre de texture proposé. . . . .	61
2.19	Estimation des paramètres de texture . . . . .	63
2.20	Filtres de Gabor . . . . .	63
2.21	Énergies de Gabor calculées à la sortie de filtres orientés (vidéo inversée). . . . .	65
2.22	Énergies de Gabor calculées à la sortie de filtres orientés (vidéo inversée). . . . .	66
2.23	Paramètres estimés dans la direction N, i.e. $\mathbf{d} = (\mathbf{0}, -\mathbf{1})$ (vidéo inversée). . . . .	68
<b>Chapitre 3.</b>		<b>71</b>
3.1	(a): Image panchromatique SPOT3 (b): MRI d'un cerveau (echo T1 d'un scanner sigma 1.5 Tesla, GE). . . . .	82
3.2	Exemples d'images des degrés d'appartenance à 3 classes : image SPOT3. . . . .	83
3.3	Exemples d'images des degrés d'appartenance à 3 classes : IRM d'un cerveau. . . . .	83
3.4	Classification (a),(c): avec KM (b),(d): avec FCME. . . . .	84
3.5	(a), (b): Simulations SPOT5. (c) et (d): Images du paramètre de texture estimé. (e) et (f) histogrammes des images (c) et (d). . . . .	85
3.6	Images des degrés d'appartenance à chacune des 2 classes (FCME). . . . .	86
3.7	Classification (a),(b): avec KM (c),(d): avec FCME. . . . .	86
3.8	Segmentation obtenue par l'ICM (contours blancs) et marqueurs (régions grises). . . . .	91

3.9	Canet en Roussillon : simulation SPOT5 à 5m. . . . .	97
3.10	Toreilles : simulation SPOT5 à 5m. . . . .	97
3.11	Eppeville : simulation SPOT5 à 5m. . . . .	98
3.12	Canet en Roussillon : simulation SPOT5 à 5m. . . . .	99
3.13	Toreilles : simulation SPOT5 à 5m. . . . .	100
3.14	Toulouse : simulation SPOT5 à 2m50. . . . .	101
3.15	La Roche sur Yon : image SPOT3 à 10m. . . . .	102
3.16	Cayenne : image SPOT3 à 10 m. . . . .	103
3.17	Macapá (Brésil) : images ERS1 à 12m50 (8 bits). . . . .	104
3.18	Projection des limites des zones urbaines obtenues par photo- interprétation d'images radar sur la carte topographique (1:100 000, 1995) fournie par l'Institut Brésilien de Géographie et de Statistique.	105
3.19	Belém (Brésil) : image ERS1 à 12m50 (8 bits). . . . .	106
3.20	Macapá (Brésil) : image ERS1 à 12m50 (16 bits). . . . .	107
3.21	Aix en Provence : image ERS1 à 12m50 (16 bits). . . . .	108
3.22	Schéma global. . . . .	109
<b>Chapitre 4.</b>		<b>111</b>
4.1	Transformations de base (dilatation, érosion, ouverture, fermeture) par un disque de diamètre 17. . . . .	116
4.2	Distance géodésique versus distance euclidienne. . . . .	120
4.3	Élément structurant unité défini en 4-connexité et 8-connexité. . . . .	122
4.4	Sections supérieures. . . . .	122
4.5	Ouverture et ouverture par reconstruction de l'image présentée 4.1 (a) par un disque de diamètre 17 . . . . .	124
4.6	Ouverture (CPU=0s57) et ouverture par reconstruction (CPU=3s89) par un disque de diamètre 7 de l'image d'Aix en Provence. . . . .	125
4.7	Ouverture par reconstruction (CPU=8s47) par un disque de diamètre 7 de l'image de La Roche sur Yon. . . . .	126
4.8	Simulation SPOT5 à 5m . . . . .	130

4.9	Exemple de granulométrie sur l'image (Figure 4.6 (a)), premières itérations : $r \in \{1, 2, 3, 4, 5, 6\}$ . . . . .	139
4.10	(a)(b): Aix en Provence, (c)(d): Marseille . . . . .	140
4.11	Classes sélectionnées . . . . .	141
4.12	Segmentation finale en 4 classes. . . . .	142

# Tableaux

<b>Chapitre 1.</b>	<b>5</b>
1.1 Principaux satellites à haute résolution. . . . .	7
<b>Chapitre 2.</b>	<b>25</b>
2.1 Temps de calcul pour chacune des 3 méthodes sur station SUN ultra1 (124MB, 167 MHz). . . . .	67
<b>Chapitre 3.</b>	<b>71</b>
3.1 Paramètres de l'algorithme FCME. . . . .	81
3.2 Valeurs des paramètres et temps de calcul sur station SUN ultra1 (124MB, 167 MHz). . . . .	96
<b>Chapitre 4.</b>	<b>111</b>
4.1 Transformations de base dans le cas binaire. . . . .	117
4.2 Transformations de base dans le cas des images en niveaux de gris. . .	118
4.3 Valeurs des paramètres et temps de calcul sur station SUN ultra1 (124MB, 167 MHz). . . . .	136



# Introduction

**D**ans le contexte de l'expansion du marché des satellites civils à haute résolution, nous nous sommes intéressés à l'utilisation des images satellitaires pour l'extraction d'un masque urbain et l'analyse intra-urbaine. Le travail de thèse, réalisé dans le cadre d'une collaboration CNES-INRIA, est présenté de manière succincte dans ce chapitre.

Le sujet de cette thèse a été proposé conjointement par l'INRIA et par le CNES. Les objectifs principaux étaient de développer de nouveaux modèles, de nouveaux algorithmes de classification et de segmentation en vue d'extraire et d'analyser le paysage urbain. Cette analyse se plaçait dans le contexte du développement des satellites civils à haute résolution.

Dans le but d'analyser le paysage urbain nous avons procédé par étapes. La première a consisté à extraire un masque puis la seconde à caractériser la zone intra-urbaine.

Nous avons utilisé comme données des images satellitales à haute résolution (simulations SPOT5 [11]) mais également des images prises par SPOT3. Par ailleurs nous avons également extrait des masques urbains sur des images radar prises par ERS1.

Le manuscrit se divise en quatre chapitres.

Dans le premier chapitre, nous présentons le contexte dans lequel ce travail a été mené, les applications possibles, les types de données utilisées, ainsi que les méthodes appliquées. Nous présentons notamment le contexte général qui se caractérise par une croissance future du marché civil des satellites à haute résolution, liée essentiellement aux lancements programmés dans les années à venir de plusieurs satellites de résolution métrique. À cela s'ajoute une extension des domaines d'application en urbain. Quelques exemples de projets européens sont présentés. Par la suite, nous précisons les conditions dans lesquelles nous avons travaillé et notamment les données mises à notre disposition. A ce sujet, un intérêt particulier a été accordé aux données simulées SPOT5. Dans cette optique, nous décrivons brièvement les conditions d'obtention de ces simulations à partir d'images aériennes.

Au terme de cette première partie, nous décrivons le cadre méthodologique qui a guidé notre travail : la modélisation statistique, les champs de Gibbs, les champs de Markov, l'inférence bayésienne. Ainsi, les définitions générales, les notations utilisées à diverses reprises dans ce manuscrit sont données une seule fois. Finalement, nous présentons un état de l'art des techniques d'analyse des textures.

Les deux chapitres suivants sont consacrés respectivement aux méthodes d'analyse de texture puis de classification/segmentation spécialement développées dans le but d'extraire un masque urbain .

Dans un premier temps, un état de l'art des méthodes d'extraction d'un masque urbain ainsi que quelques applications sont présentés.

Nous nous focalisons, par la suite, sur deux méthodes d'analyse isotrope de la texture, l'une fondée sur un modèle markovien isotrope 2D, l'autre fondée sur les **matrices aura** dérivées des matrices de cooccurrence. Nous décrivons ensuite le modèle markovien fondé sur des chaînes que nous avons proposé. Finalement, les résultats obtenus sont comparés avec les analyses classiques par matrices de cooccurrence et par bancs de filtres de Gabor.

Dans le troisième chapitre, nous nous intéressons au problème de la classification et de la segmentation de l'image du paramètre extrait à l'étape précédente. Dans cette optique, nous proposons un algorithme de classification floue qui ne requiert pas la connaissance du nombre de nuages de points (ou *clusters* en anglais). Cette approche consiste à minimiser un critère qui tient compte à la fois de la distance intra-nuage (i.e. la distance entre les points appartenant à un même nuage) et d'un terme d'entropie. L'algorithme estime en même temps le nombre de nuages de points et la matrice de partition. Les paramètres ainsi estimés sont finalement introduits dans un modèle markovien de segmentation qui utilise le modèle de Potts comme modèle de régularisation.

Des résultats expérimentaux, sur des types de données différents à la fois en termes de capteur (optique, radar) et de résolution (2m50, 5m, 10m, 12m50), sont analysés.

Dans le quatrième chapitre, nous présentons la méthode utilisée afin d'analyser le paysage urbain, i.e. d'extraire plusieurs classes dans les zones urbaines selon certains critères. Auparavant, nous rappelons quelques études réalisées précédemment ainsi que quelques applications.

L'approche est basée sur une transformation de la morphologie mathématique : l'ouverture par reconstruction. Elle est utilisée comme transformation de base d'une granulométrie. Cette analyse nous permet d'extraire une information caractéristique de la taille, de la forme des objets. Elle est finalement intégrée dans un modèle markovien de segmentation. Nous présentons quelques résultats expérimentaux réalisés essentiellement sur des simulations SPOT5.

La conclusion nous permet de mettre en évidence les apports de ce travail tout en soulignant les limites ainsi que les points à approfondir.

## DANS CE CHAPITRE:

1.1	Contexte Général . . . . .	6
1.2	Contexte Particulier . . . . .	9
1.2.1	Données . . . . .	10
1.2.2	Notations et Définitions . . . . .	12
1.2.2.1	Inférence bayésienne . . . . .	13
1.2.2.2	Distribution a priori et a posteriori . . . . .	13
1.2.2.3	Risque et estimateur bayésien . . . . .	13
1.2.2.4	Champs de Markov . . . . .	14
1.2.2.5	Champs de Gibbs . . . . .	15
1.2.3	Analyse de texture : état de l'art . . . . .	16
1.2.3.1	Méthodes statistiques et spectrales . . . . .	17
1.2.3.2	Méthodes géométrique . . . . .	17
1.2.3.3	Méthodes de filtrage . . . . .	18
1.2.3.4	Méthodes fondées sur un modèle . . . . .	20
1.2.3.5	Méthodes mixtes . . . . .	23
1.2.3.6	Conclusion . . . . .	23

# 1.

## Contexte - Objectifs - Méthodes

Avant de rentrer dans le vif du sujet, nous nous proposons de replacer la problématique de cette thèse dans un cadre plus général. Dans cette optique, nous présentons, dans un marché des satellites civils à haute résolution en pleine expansion, les nombreuses applications possibles de l'analyse urbaine et quelques projets mis en oeuvre par divers organismes européens. Puis nous décrivons plus particulièrement notre cadre de travail, i.e. les conditions dans lesquelles nous avons effectué notre travail de recherche à la fois en termes de méthodologie et de type de données utilisées. Finalement, un état de l'art des méthodes d'analyse des textures est présenté.

## 1.1 Contexte Général

Nous replaçons ici le sujet de recherche dans un cadre plus général qui tient compte à la fois de l'évolution des données à traiter (vers la haute résolution) et également des applications qui peuvent justifier une telle étude.

Les grands programmes (de type SPOT) vont permettre d'accéder dans les années à venir des images à haute résolution. Par ailleurs, depuis mars 1994 et la décision du président américain d'autoriser le lancement des satellites à haute résolution (1 à 3 m) à des fins commerciales et civiles, de nombreux projets ont vu le jour (voir le tableau 1.1).

Avec l'émergence de nouveaux domaines d'applications notamment dans les télécommunications et l'amélioration de la résolution des images satellitaires, l'intérêt pour l'analyse du paysage urbain par télédétection satellitale s'est accru.

Certes depuis de nombreuses années des organismes comme l'ORSTOM [5] ou l'IGN [62, 74], utilisent les images satellitaires pour des applications qui sont tournées vers l'étude de l'évolution, l'expansion des zones urbaines notamment dans les pays émergents (étude sur QUITO de l'ORSTOM).

L'avènement de la moyenne résolution ( $\approx 10m$ ) puis de la haute résolution ( $\approx 2m$ , SPOT5, Quickbird) laisse entrevoir des applications variées. Soulignons d'ores et déjà que cette avancée en matière de résolution de l'image n'ira pas sans poser de nouveaux problèmes au niveau du traitement, et de l'analyse de l'image. Ainsi, pour un thème donné, le bruit au sens du thème sera plus important : cas des ombres portées, des arbres en bordures des rues, des voitures  $\dots$ . En effet, lorsque l'on cherche à extraire un réseau routier par exemple, les véhicules éventuellement visibles sur l'image perturbent la bonne détection du thème "route" en introduisant des contours supplémentaires. Dans ce sens, les véhicules sont considérés comme du bruit pour le thème "route" que l'on veut extraire.

Ces dernières années, plusieurs projets européens ont été lancés avec l'objectif d'utiliser l'imagerie satellitale à haute et à très haute résolution pour analyser le paysage urbain. Ainsi, le projet EUROSTAT a été initié par SCOT Conseil et s'est achevé en 1997. Plus récemment, le projet ATLAS qui propose une étude sur les villes de Berlin et d'Athènes a été lancé. Par ailleurs, certaines communautés urbaines telles que Toulouse se sont intéressées aux potentialités de l'imagerie satellitale pour la gestion de la ville.

Achévé en mai 1997, EUROSTAT<sup>1</sup> [77] avait pour but essentiel de définir et

---

<sup>1</sup><http://www.scot-sa.com/urbain/euro1.htm>

Nom	Résolutions		Lancement prévu	Instigateurs
	Panchromatique	Multispectral		
SPOT5	2m50-5m	10m	2002	France-Suède-Belgique
IRS-1D	5m80	23m-70m	29.09.97	Inde
ADEOS	8m	16m	1996 perdu 1997	Japon
KVR-1000	2-3m	-	-	Russie
ALOS	2m50	10m	2002	Japon
Earlybird	3m	-	explosé 28.12.97	Earthwatch Consortium
Quickbird	1m	4m	1999	Earthwatch Consortium
IKONOS 1	1m	4m	27.04.99 contact perdu	Space Imaging EOSAT
EROS A-EROS B	1m50	-	?	Israël
OrbView-3	1m	4m	2000	Orbital Imaging Corporation

**Tableau 1.1:** *Principaux satellites à haute résolution.*

tester des méthodes pour délimiter les agglomérations européennes au moyen de la télédétection, d'informations géographiques, de données cadastrales. D'autre part, les participants (SCOT Conseil, CESD-Communautaire et 6 villes européennes) ont défini au niveau européen, en collaboration avec des Instituts Nationaux de la Statistique, une nouvelle nomenclature à 4 niveaux (CLUSTER) afin de mieux répondre aux attentes des statisticiens.

Quant au projet ATLAS<sup>2</sup>, il est plus directement lié à l'arrivée sur le marché civil des satellites à haute résolution (<3m). Il fait partie de l'un des 5 projets lancés par le CEO (*Center for Earth Observation*) afin de favoriser l'utilisation des données produites par les satellites d'observation de la terre. L'étude porte plus précisément sur les villes de Berlin et d'Athènes. Plusieurs organismes européens (GEOAPIKONISIS L.T.D, GAF mbh, SCOT Conseil, VRS GmbH, The Technical university of Berlin) ont mis en commun leurs compétences (pré-traitement des données, collecte d'autres sources d'information, classification automatique des données satellitales, validation des résultats à partir de vérités terrain) pour mener à bien ce projet. Les données utilisées sont des images satellitales de IRS-1C (satellite indien) et KVR-1000 (caméra sur satellite russe KOSMOS) et des compositions (scènes multispectrales de IRS-1C + scène panchromatique de IRS-1C ou scènes multispectrales de IRS-1C + KVR-1000). Par ailleurs, des données exogènes peuvent être intégrées (e.g. cartes topographiques, des cartes du biotope ...) dans le but d'améliorer la résolution, de prendre en compte des informations plus détaillées, de mettre en évidence les changements détectés pour l'utilisateur final. Les paramètres utilisés pour la classification de l'image satellitale font appel à des calculs de variances locales, à la théorie de l'évidence de Dempster-Shafers, aux réseaux de neurones.

L'extraction des thèmes bâti/non bâti est par ailleurs utilisée en situation réelle pour gérer l'environnement. Ainsi, les scientifiques du Parc régional des Vosges du nord ont pu recouper cette information (bâti/non bâti) avec une carte d'occupation du sol afin d'aider les autorités locales à gérer au mieux l'occupation du sol et notamment optimiser les constructions sur les zones industrielles.

En ce qui concerne la gestion des territoires des études ont été menées en 1996 conjointement par SCOT Conseil et la faculté des sciences et techniques de Tanger<sup>3</sup> en vue d'étudier la pression urbaine (ville de Tanger) sur les milieux forestiers, agricoles et sur le littoral. Les données à leur disposition étaient des compositions (P+XS) à 10 m de résolution ainsi que des SIG (Systèmes d'Informations Géographiques). La méthode appliquée était une segmentation en zones homogènes,

---

<sup>2</sup><http://ewse.ceo.org/anonymous/construct/details.pl/674539>

<sup>3</sup><http://www.scot-sa.com/tanger/>

puis une classification par nuées dynamiques en 30 classes, suivi d'un regroupement par photo-interprétation en fonction de la réalité terrain.

De même en 1991-1992, dans le cadre du projet EPUTAT, SCOT Conseil, Géoimage et le Centre National de Télédétection (CNT) tunisien ont étudié à partir d'images SPOT et Landsat la pression urbaine sur les zones agricoles de la ville de Tunis.

Un des domaines d'application prometteurs est certainement celui des télécommunications. En effet, avant de mettre en place un nouveau réseau de radiotéléphonie cellulaire il est nécessaire de créer des modèles de prédiction de propagation des ondes. Pour ce faire, les maîtres d'œuvre ont notamment besoin de cartes d'occupation du sol appelées *clutter map*. Dans le cas de l'étude conjointe réalisée sur la ville de Londres par le CNES, la société ISTAR (FRANCE) et VODAFONE (UK) en 1994, les thèmes urbains ont été analysés par photo-interprétation. Le but de cette étude était de mettre en évidence l'intérêt de l'imagerie satellitale par rapport à l'imagerie aérienne dans ce domaine.

Les quelques exemples d'applications et de projets décrits précédemment donnent une idée de l'utilité croissante d'une analyse du paysage urbain. Il faut noter que dans la plupart des cas d'utilisation en situation réelle (e.g. gestion du parc régional des Vosges, réalisation de *clutter map* pour le compte de sociétés mettant en place des réseaux de téléphonie mobile) l'analyse du paysage urbain est réalisée par photo-interprétation. Ces considérations mettent en avant la nécessité d'une automatisation, partielle ou totale, des processus afin d'aider le photo-interprète. D'autres exemples plus directement liés soit à l'extraction d'un masque urbain (2.2), soit à l'analyse intra-urbaine (4.2) seront présentés dans la suite du manuscrit.

## 1.2 Contexte Particulier

Compte tenu de l'accroissement de la demande en ce qui concerne l'analyse des zones urbaines notamment dû à l'explosion des télécommunications, il nous paraît intéressant d'automatiser les processus. L'automatisation du traitement des images a pour but essentiel de minimiser les décisions subjectives (la variabilité d'interprétation entre deux opérateurs humains est de l'ordre de 10% pour une classification à partir d'une même image), en vue notamment de la mise à jour des cartes d'occupation du sol. Par ailleurs, elle permettrait de traiter des zones du monde peu documentées du fait notamment de leur inaccessibilité. Enfin, cette automatisation a également pour but de diminuer les temps de traitement. En effet, le géographe met par exemple une année pour réaliser une cartographie de

villes brésiliennes telles que Macapá ou Belém du fait notamment de la grande hétérogénéité des constructions. En outre, ces cartes ne seront probablement plus valables au bout de 2 ans du fait de l'évolution très rapide de ce type de ville. La ville de Macapá voit, par exemple, sa population augmenter d'environ 10% par an.

Notre approche a consisté à développer des modèles statistiques dans le but d'analyser la texture de l'image, de la classifier, de la segmenter de la manière la plus automatique possible. Par ailleurs, les algorithmes de classification et de segmentation proposés sont applicables de manière beaucoup plus générale à de nombreux problèmes de traitement d'image. L'analyse urbaine n'étant qu'une des applications possibles.

L'image satellitale, et plus précisément le thème urbain, a guidé notre recherche dans la toute première étape de l'analyse. En effet, l'observation visuelle de l'image ainsi que la mise en évidence des écueils des méthodes classiques (e.g. morphologie mathématique [108]) nous ont permis de définir un nouveau modèle utilisé afin d'analyser la texture urbaine.

Par rapport aux projets (de type EUROSTAT) présentés précédemment, nous nous plaçons dans une démarche différente. En effet, nous travaillons dans la mesure du possible dans un cadre non supervisé. Très souvent ces équipes font intervenir des photo-interprètes dans la chaîne de décision. Enfin, nous ne disposons pas de plusieurs types de données (SIG, données cadastrales, vérités de terrain ...). Nous n'utilisons que les images satellitales optiques et radars mises à notre disposition par divers organismes (CNES, ESA). Enfin, nous ne travaillons pas sur une ville particulière comme cela est souvent le cas dans les projets type ATLAS décrit précédemment.

### 1.2.1 Données

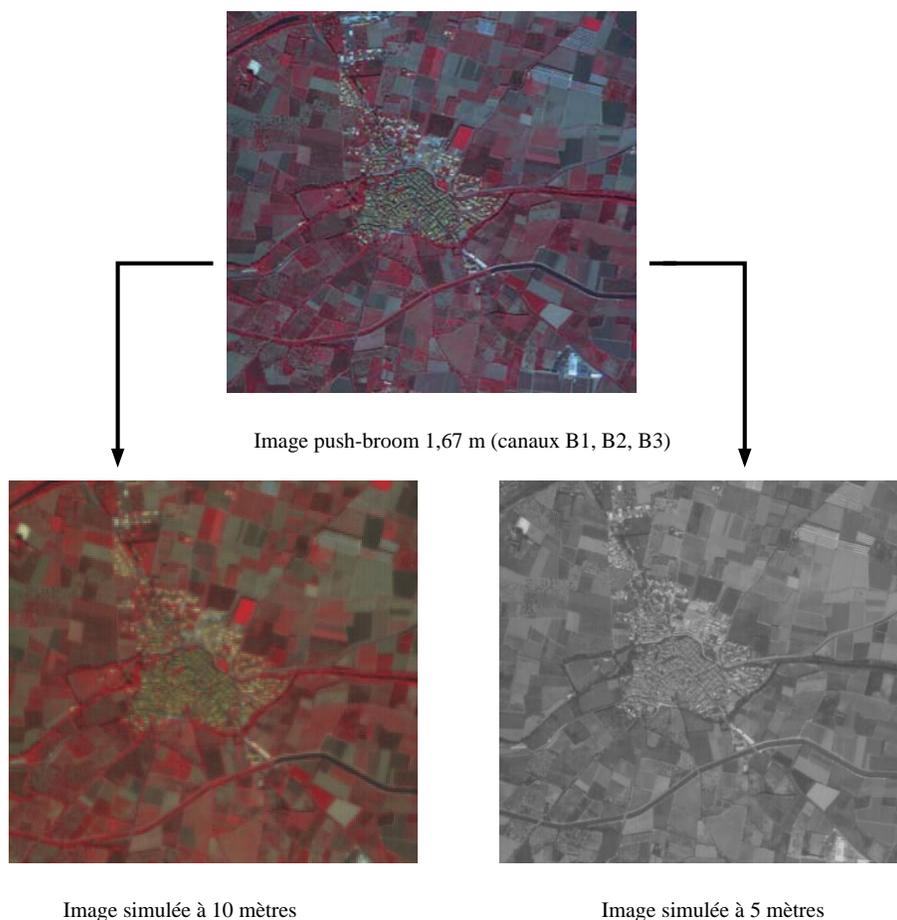
La caractérisation du thème urbain fait davantage appel aux images aériennes et aux enquêtes sur le terrain qu'aux images satellitales. Cependant, avec l'arrivée des satellites civils à haute résolution, on s'attend à un développement du marché. Nous avons travaillé sur des données satellitales optiques SPOT fournies par le CNES. Plus précisément, nous nous sommes intéressés aux potentialités de SPOT5 [11] sans que cela soit limitatif. En effet, il paraissait également intéressant de voir les limites de notre approche par rapport à la résolution de l'image. C'est pourquoi nous avons également utilisé des images SPOT3. Par ailleurs, nous avons également eu l'opportunité de tester nos algorithmes sur des images radar ERS1 fournies par l'ESA. Les résultats sont prometteurs.

Le programme SPOT (Satellite Pour l'Observation de la Terre) décidé en 1978 par la France, la Suède et la Belgique est opérationnel depuis février 1986 avec le lancement de SPOT1. Les trois premiers satellites possèdent les mêmes caractéristiques (notamment 3 bandes spectrales avec une résolution de 20m, un mode panchromatique avec une résolution de 10m). A l'heure actuelle seuls SPOT1 et 2 fonctionnent encore. En orbite depuis mars 1998, SPOT4 offre quant à lui une bande spectrale supplémentaire dans le moyen infra-rouge. SPOT5 présentera, entre autre, l'avantage d'avoir une meilleure résolution au sol (2m50 et 5m en mode panchromatique et 10m en mode multispectral). Ces nouveaux capteurs permettront ils de mieux répondre aux demandes du marché notamment en ce qui concerne l'analyse urbaine ? Des études préliminaires sur des images simulées sont effectuées afin d'essayer de répondre à cette question. Tous ces satellites ont été, sont ou seront placés sur la même orbite circulaire, héliosynchrone et phasée par rapport à la terre à 830 km d'altitude.

Le satellite SPOT5 ne sera en orbite qu'au début des années 2000. En attendant, le CNES fournit des images simulées aux divers organismes intéressés. Les simulations SPOT5 [96] sont obtenues à partir du traitement au sol d'images obtenues avec un radiomètre aéroporté (système RAMI : Radiomètre Aéroporté Multispectral Imageur). L'instrument utilise un type de balayage appelé PUSH-BROOM identique à celui du système SPOT. Une ligne de l'image est obtenue à partir d'une barrette CCD (*Charge Coupled Device*) qui délivre un signal électrique proportionnel à l'éclairement provenant d'un pixel au sol. L'image totale est obtenue par déplacement de l'avion.

La bande spectrale panchromatique est obtenue à partir des canaux 1 et 2. Les autres bandes spectrales sont directement obtenues à partir de celles du PUSH-BROOM. Ainsi, nous obtenons 4 canaux de résolution spatiale 1m67. Le passage aux résolutions de 5m et 10m se fait par un filtrage suivi d'un sous échantillonnage. Les paramètres (facteur de sous-échantillonnage 3 et 6 et taille du masque 15 et 31) dépendent de la résolution voulue. Le filtre est défini à partir de la transformée de Fourier inverse de la FTM (Fonction de Transfert de Modulation) de SPOT5. Les simulations SPOT5 à 2m50 sont obtenues à partir de 2 images à 5m obtenues avec un capteur matriciel et décalées selon le processus appelé SUPERMODE (échantillonnage en quinconce).

Le satellite ERS1 a été lancé en juillet 1991. Il est composé de plusieurs instruments dont un radar à synthèse d'ouverture (RSO). Les images radar prises par ERS1 et fournies par L'Agence Spatiale Européenne (ESA) sont codées sur 16 bits



**Figure 1.1:** *Exemple de simulations SPOT5 : Toreilles.*

et ont une résolution de 12m50.

### 1.2.2 Notations et Définitions

Les méthodes appliquées lors de ce travail sont pour l'essentiel fondées sur une analyse probabiliste de l'image et une approche bayésienne.

La modélisation markovienne est ainsi utilisée à plusieurs reprises notamment lors de l'étape de caractérisation de la texture urbaine ainsi que lors des étapes de segmentation. C'est pourquoi, dans le paragraphe suivant, nous donnons quelques notations régulièrement employées dans ce manuscrit ainsi que quelques définitions [128].

### 1.2.2.1 Inférence bayésienne

Dans tous les problèmes classiques de traitement d'image (restauration, classification, détection de contours  $\dots$ ) nous avons accès à des observations  $y$  qui sont fonction de l'image originale  $x$ .

$y$  est souvent une version dégradée de  $x$ , e.g. une version bruitée :

$$y = x + \nu, \text{ où } \nu \text{ est une famille de variables aléatoires} \quad (1.1)$$

L'ensemble des données est noté  $\mathbf{Y}$  et l'ensemble des images  $\mathbf{X}$ .

L'inférence bayésienne consiste à estimer  $x$  à partir des observations  $y$ .

### 1.2.2.2 Distribution a priori et a posteriori

$P(Y = y|X = x)$  est la probabilité d'observer  $Y = y$  sachant que  $x$  est l'image originale.

Dans la suite, nous noterons  $P(X)$  pour  $P(X = x)$ ,  $X$  étant la variable aléatoire et  $x \in \mathbf{X}$  une réalisation.

La probabilité a priori notée  $\pi(X)$  ne dépend pas des observations et peut donc être définie avant l'observation des données, d'où le terme probabilité a priori.  $P(X \in \mathbf{X}, Y \in \mathbf{Y}) = \pi(X)P(Y|X)$  représente la probabilité que  $X$  soit l'image originale et  $Y$  l'image observée.

Supposons maintenant que les données sont observées. La probabilité a posteriori de  $X$  sachant  $Y$ , notée  $P(X|Y)$ , peut alors être interprétée comme un ajustement de  $\pi(X)$  aux données [128]. Nous pouvons réécrire  $P(X|Y)$  de la manière suivante :

$$P(X|Y) = \frac{\pi(X)P(Y|X)}{\sum_{Z \in \mathbf{X}} P(Y, Z)} = \frac{\pi(X)P(Y|X)}{P(Y)} \quad (1.2)$$

### 1.2.2.3 Risque et estimateur bayésien

Les problèmes de restauration, classification, segmentation ( $\dots$ ) consistent à sélectionner parmi un ensemble d'images la "meilleure" image  $\hat{X}$ , au sens d'un certain critère, qui soit "compatible" avec les données observées  $Y$ . Ce problème est

indéterminé dans la mesure plusieurs images  $\hat{X}$  peuvent convenir. L'idée est de contraindre le système afin de réduire le nombre de solutions. Ainsi, l'analyse bayésienne permet de tenir compte à la fois de la correspondance entre  $\hat{X}$  et  $Y$  et de critères de qualité comme par exemple l'homogénéité des régions.

Dans le cadre de la théorie de l'estimation, les estimateurs sont définis à partir d'une fonction de coût  $L$ . Cette fonction mesure la distance entre deux configurations. Nous notons ainsi  $L(X', X)$  la distance entre les configurations  $X'$  et  $X$ . On définit alors le risque bayésien  $R$  comme étant l'espérance de  $L(X', X)$ , i.e. :

$$R(X') = E[L(X', X)|Y] = \int_{\mathbf{X}} L(X', X)P(X|Y)dX \quad (1.3)$$

Un estimateur bayésien  $\hat{X}$  est un estimateur qui minimise ce risque.

On notera :  $\hat{X} = \underset{X'}{\operatorname{argmin}}[R(X')] = \underset{X'}{\operatorname{argmin}}[E[L(X', X)|Y]]$

La fonction  $L$  la plus simple est définie comme suit :

$$L(X', X) = \begin{cases} 0 & \text{si } X' = X \\ 1 & \text{si } X' \neq X \end{cases} \quad (1.4)$$

L'estimateur  $\hat{X}$  qui minimise le risque bayésien correspondant est appelé estimateur du Maximum A Posteriori (MAP) car il maximise la probabilité a posteriori de  $X$  sachant  $Y$  ( $\hat{X}_{MAP} = \underset{X}{\operatorname{argmax}}[P(X|Y)]$ ).

Un estimateur bayésien très utilisé notamment en classification est l'estimateur du MPM (*Marginal Posterior Mode*).

De nombreux autres estimateurs bayésiens peuvent ainsi être définis. Ces estimateurs vont dépendre de la fonction  $L$  mais également du modèle a priori, i.e. du choix de  $\pi(X)$ .

#### 1.2.2.4 Champs de Markov

Les champs de Markov sont largement utilisés en analyse d'image [6, 34, 43]. Nous les avons utilisés à deux reprises. Dans une première étape, nous modélisons la texture de l'image par huit champs markoviens gaussiens fondés sur des chaînes (paragraphe 2.4), ce qui nous permet de définir un paramètre caractéristique de la texture urbaine. La modélisation markovienne utilisée pour définir un modèle a priori, nous permet également de segmenter les images (paragraphe 3.2).

Les champs de Markov [43, 128] permettent de modéliser des propriétés globales en utilisant des contraintes locales.

Nous notons  $S$  ( $\text{card}(S) = |S| = N$ ) le réseau des sites de l'image.

Soit  $V = \{V(s) | s \in S\}$  le système de voisinage ( $V(s)$  est le voisinage du site  $s$ ).

$V$  est un système de voisinage si :

(i)  $s \notin V(s)$ .

(ii)  $s \in V(t) \Leftrightarrow t \in V(s)$ .

Nous notons  $X = \{X_1, \dots, X_N\}$  une famille de variables aléatoires définies sur  $S$ . Chaque variable aléatoire  $X_i$  est à valeur dans  $\Omega$ , l'espace des états<sup>4</sup>. Nous notons  $X_i = x_i$ ,  $x_i \in \Omega$ , pour signifier que la variable aléatoire  $X_i$  prend la valeur  $x_i$ . De même, nous notons  $X = x$ ,  $x \in \Omega^N$ , pour signifier l'événement suivant :  $(X_1 = x_1, \dots, X_N = x_n)$ .

$x$ , appelé configuration, correspond à une réalisation du champ aléatoire  $X$ . Cette configuration a une certaine probabilité que nous notons  $P(X = x)$ .

$(X, \Omega^N, P)$  est appelé champ de Markov par rapport au système de voisinage  $V$  s'il vérifie les 2 propriétés suivantes :

- $\forall x \in \Omega^N, P(X = x) > 0$  (positivité)
- $\forall s \in S$  et  $\forall x \in \Omega^N$ ,  
 $P(X_s = x_s | X_r = x_r, r \in S \setminus \{s\}) = P(X_s = x_s | X_r = x_r, r \in V(s))$  (propriété markovienne)

### 1.2.2.5 Champs de Gibbs

En utilisant les mêmes notations que précédemment, nous définissons les champs de Gibbs.

$(X, \Omega^N, P)$  est un champ de Gibbs induit par la fonction d'énergie  $\mathcal{U}$  si  $P(X)$  est de la forme :

$$P(X) = \frac{\exp(-\mathcal{U}(X))}{\sum_{X \in \Omega^N} \exp(-\mathcal{U}(X))} \quad (1.5)$$

Parmi tous les champs de Gibbs, nous nous intéressons plus particulièrement aux champs de portée bornée, i.e. dont les interactions sont locales.

---

<sup>4</sup> $\Omega$  peut être discret, ou continu.

Le théorème de Hammersley-Clifford [14, 43] lie la propriété de localité des champs de Markov (propriété markovienne) et la propriété de globalité des champs de Gibbs (distribution de Gibbs) [14, 79].

**Théorème 1.2.1** *Si  $(X, \Omega^N, P)$  est un champ de Markov par rapport au système de voisinage  $V$  alors sa distribution  $P(X)$  est une distribution de Gibbs de portée bornée, i.e. :*

$$P(X) = \frac{\exp(-\mathcal{U}(X))}{Z}$$

$Z = \sum_{X \in \Omega^N} \exp(-\mathcal{U}(X))$  est une constante de normalisation appelée fonction de partition.

$\mathcal{U}(X)$  est une fonction d'énergie définie comme suit :

$$\mathcal{U}(X) = \sum_{c \in \mathcal{C}} \mathbf{V}_c(X_s, s \in c)$$

$c$  est une clique, i.e. un ensemble de sites voisins deux à deux ou un singleton.

$\mathcal{C}$  est l'ensemble des cliques liées au système de voisinage considéré.

$\mathbf{V}_c(\cdot)$  est une fonction de potentiel.

L'énergie globale  $\mathcal{U}(X)$  s'écrit donc comme une somme d'énergies locales (ou somme de potentiels).

### 1.2.3 Analyse de texture : état de l'art

L'analyse des zones urbaines que nous proposons est fondée sur une analyse de la texture de l'image. C'est pourquoi nous présentons dans ce qui suit un état de l'art dans ce domaine.

De manière générale, on distingue 5 types de méthodes : les méthodes structurelles-géométriques (davantage réservées à l'étude des macro-textures), les méthodes statistiques, les méthodes de filtrage, les méthodes fondées sur un modèle, et les méthodes mixtes.

Il est très difficile de définir la notion de texture de façon précise. Tout au plus pouvons nous donner quelques caractéristiques permettant de cerner cette notion. Classiquement on distingue les micro-textures (e.g. sable) dans lesquelles nous

ne distinguons visuellement aucune figure géométrique particulière mais qui sont caractérisées par une certaine homogénéité statistique, des macro-textures (e.g. tissu écossais) dans lesquelles une même figure géométrique semble se répéter plus ou moins régulièrement. Plusieurs articles proposent une très bonne synthèse des méthodes existantes [55, 118].

Dans la suite de ce manuscrit nous décrivons quelques approches à titre d'exemple et ce pour chacune des 5 catégories. Notons que la plupart des méthodes décrites ci-dessous permettent une analyse anisotrope de la texture.

### 1.2.3.1 Méthodes statistiques et spectrales

Dans [55], Haralick extrait une série de paramètres (e.g. l'énergie, le contraste, l'entropie, la corrélation  $\dots$ ) à partir des matrices de cooccurrence, qu'il utilise pour classifier des images LANDSAT. Dans [22], les auteurs proposent deux nouveaux paramètres extraits des matrices de cooccurrence (*cluster shade* et *cluster prominence*) davantage adaptés à la segmentation de zones urbaines.

Dans [118], la texture est analysée à partir du spectre de puissance de l'image dont on extrait des paramètres. Le plan fréquentiel dans lequel est représenté ce spectre est partitionné en anneaux centrés à l'origine ainsi qu'en secteurs angulaires. L'énergie est alors calculée dans chaque anneau, puis dans chaque secteur donnant des paramètres caractéristiques respectivement de la finesse de la texture ainsi que de sa directionnalité. Ces méthodes sont davantage adaptées à l'analyse des micro-textures.

### 1.2.3.2 Méthodes géométrique

Les méthodes géométriques considèrent la texture comme un ensemble d'éléments de base appelés textons [64, 117] dont la disposition est organisée grâce à des règles de placement. Zucker [134] a développé un modèle qui lui permet à partir d'une texture idéale de synthétiser une texture observée qui correspond à une version distordue de la première. La texture idéale est composée de primitives régulières (ex. le carré) disposées régulièrement. Des transformations déterministes ou stochastiques sont appliquées au graphe d'étiquettes constitué à partir de la texture idéale. La version distordue est alors obtenue par transformation inverse. Dans [40], Fu propose de représenter la texture à l'aide d'une grammaire d'arbre (inférence grammaticale). Cette approche s'explique aisément si l'on divise l'image de texture en fenêtres. Chacune d'elle représente une primitive (premier niveau) qui est elle-même divisée en fenêtres jusqu'à obtenir l'élément de base qui va constituer

la racine de l'arbre. Ainsi l'ensemble des primitives va définir une grammaire à laquelle on va associer des règles de placement. A ces dernières sont parfois associées des probabilités qui vont pouvoir influencer par exemple sur la luminosité de la texture. La méthode proposée permet non seulement la synthèse de texture mais également la segmentation d'images. La notion de hiérarchie est ici présente et on la retrouvera avec les approches multirésolution (filtres de Gabor, ondelettes  $\dots$ ).

Enfin, les modèles de mosaïques [1] permettent de générer des textures à partir de processus aléatoires. Les figures de base sont placées de manière aléatoire dans l'image. Plusieurs familles de modèles sont présentées. La première représente des modèles ayant une structure en cellules. Le plan étant divisé en cellules, on assigne de manière indépendante une couleur à chaque cellule parmi un ensemble fixé de  $N$  couleurs, suivant une série de probabilités  $p_1, \dots, p_N$ . Les modèles diffèrent par la manière dont le plan est divisé (modèle d'échiquier, modèle hexagonal  $\dots$ ). Une seconde famille est constituée des modèles "couvrants". Il s'agit d'un arrangement aléatoire de figures géométriques (chacune d'elle étant définie par son centre de gravité et ses axes principaux) dans le plan. On associe une couleur à l'intérieur de la figure et une autre à son complément dans le plan. Les centres de gravité des figures sont répartis par un processus de Poisson ponctuel. La forme et la taille des figures sont fixes. Les auteurs définissent par ailleurs des propriétés de corrélation et des propriétés géométriques qui sont caractéristiques de leurs modèles. Les méthodes géométriques sont bien adaptées à l'étude des macro-textures.

### 1.2.3.3 Méthodes de filtrage

Comme nous l'avons souligné précédemment les textures semblent bien adaptées à une analyse multi-échelle. Les études [61, 118] menées sur le système visuel humain viennent corroborer ce propos. En effet, le système visuel humain décomposerait l'image originale en versions filtrées. Ces résultats ont amené les chercheurs à s'intéresser à l'analyse de texture par filtrage multi-canal [73]. La plupart des techniques utilisent des bancs de filtres sélectifs en orientation, en fréquence et en échelle. La littérature est riche en ce qui concerne l'analyse des textures par bancs de filtres de Gabor [17, 54, 116, 126] et ondelettes [119].

Dans [63], Jain et Farrokhnia utilisent des filtres symétriques, pairs et à valeurs réelles. Les paramètres définissant chacun d'eux sont l'orientation, la fréquence radiale ainsi que la largeur des bandes pour une intensité égale à la moitié de l'intensité maximale. Cet ensemble de filtres couvre presque uniformément le domaine fréquentiel. Dans une seconde étape, les auteurs sélectionnent un sous

ensemble de filtres tel que 95% des variations d'intensité soient prises en compte. Enfin, après filtrage de l'image, ils lui appliquent une transformation non linéaire (fonction  $\tanh(\alpha t)$ ) qui peut être interprétée comme une détection de tache (*blob* en anglais). Ainsi ils obtiennent une image de paramètre caractéristique de la texture.

L'intérêt pour les ondelettes s'est largement développé ces dernières années. Ceci se vérifie également dans le domaine qui nous concerne.

Unser [119] propose une approche multi-résolution fondée sur l'analyse à l'aide d'une famille d'ondelettes redondante (*overcomplete* en anglais). Il n'a pas à sous échantillonner les images à la sortie des bancs de filtre. Sa méthode est, de ce fait, invariante par translation contrairement à celles faisant appel aux pyramides d'ondelettes, aux paquets d'ondelettes  $\dots$ . En utilisant des Bsplines, des Dsplines, l'ondelette de Battle-Lemarié ainsi que la transformée de Haar, il classe et segmente des images. De son étude il ressort que les propriétés de localisation des bancs de filtres sont plus importantes que les propriétés des fonctions de base.

Enfin, dans [84], les auteurs extraient les caractéristiques des textures avec une série d'ondelettes de Gabor. La partition est effectuée par une méthode de groupement (*clustering* en anglais) dans l'espace des caractéristiques. Ils proposent un algorithme d'apprentissage alliant les cartes de Kohonen et une méthode LVQ (*Learning Vector Quantization*). Une des applications est notamment de retrouver un motif déterminé dans une grande image.

D'autres filtres moins connus sont également utilisés. Ainsi, dans [68], les auteurs décomposent l'image à l'aide d'une série de filtres de Markov qui conservent les propriétés gaussienne et markovienne d'un processus stochastique. Cependant la sortie de ces filtres est sous échantillonnée ce qui entraîne la perte du caractère markovien. Pour y remédier ils approchent la densité du processus après le sous échantillonnage par celle du processus markovien gaussien qui l'approche au mieux au sens de la distance de Kullback-Leibler. L'algorithme ICM est utilisé pour la segmentation.

Dans le but de développer une méthode de classification invariante par rotation, Greenspan et al. [50] utilisent une pyramide laplacienne (*steerable laplacian pyramid*) afin d'obtenir des versions filtrées de l'image originale par des filtres passe-bande sélectifs du point de vue de l'orientation. Cette méthode leur permet notamment de déterminer l'angle de rotation d'une texture par rapport à une base de données.

Enfin, nous signalons également l'utilisation de la théorie de la morphologie mathématique pour l'analyse des textures [67]. Nous restons volontairement

succincts car cette approche est détaillée et utilisée dans le chapitre 4. Dans [49], les auteurs utilisent une série d'ouvertures par un disque de taille croissante (formant une granulométrie) pour analyser la texture d'images médicales et extraire le tissu interstitiel. La classification se fait par des réseaux de neurones. Dans [110], Serra utilise l'analyse granulométrique pour étudier des roches provenant du Caucase russe en termes de taille, de forme, ... afin de relier ces données avec la vitesse de propagation du son dans ces roches, leur résistance aux chocs, ce qui intéresse les géologues. Dans [106], les auteurs appliquent des filtres alternés séquentiels afin d'extraire la texture urbaine.

#### 1.2.3.4 Méthodes fondées sur un modèle

Ces méthodes consistent à choisir un modèle pour représenter le signal. Ce modèle est défini par un certain nombre de paramètres qui caractérisent la texture. Dans la littérature on retrouve très souvent les mêmes modèles. Les méthodes se distinguent davantage par la manière dont sont estimés les paramètres ainsi que par les algorithmes de segmentation, et de classification utilisés par la suite.

##### a. Modèles AR

Le modèle le plus connu et le plus utilisé est sans doute le SAR (*Simultaneous AutoRegressive Model*) qui est non causal [20]. Dans [16], les auteurs proposent un modèle de texture AR causal. L'estimation des paramètres du modèle est alors plus facile. Ce modèle est intégré dans un algorithme multi-résolution non supervisé de segmentation d'image. L'estimation des paramètres (nombre de classes et paramètres associés) est obtenue par minimisation d'un critère global fondé sur le critère AIC (*A Information Criterion*) [2]. Dans [86] les auteurs proposent un modèle SAR particulier, le modèle RISAR, qui est invariant par rotation (*Rotation Invariant Simultaneous AutoRegressive Model*). Les modèles NCAR (*Non Causal AR model*) constituent un sous groupe des GMRF (*Gaussian Markov Random Field*) décrits dans le chapitre suivant : étant donné un NCAR, on peut trouver un GMRF équivalent, l'inverse n'étant pas vrai [130].

Plus récemment, Alata et al. [3] proposent une nouvelle méthode de caractérisation des textures à l'aide de l'estimation des coefficients de réflexion 2D (fondés sur une approche multi-canal de prédiction linéaire). Un algorithme rapide d'estimation des paramètres est proposé. Leur méthode donne de meilleurs résultats

pour la segmentation d'image que celle qui utilise les paramètres AR classiques estimés à l'aide de la méthode des moindres carrés.

## b. Modèle de Wold

La décomposition de Wold fut à l'origine appliquée dans le domaine du traitement du signal 1D et plus particulièrement du traitement de la parole. Plus récemment cette méthode a été étendue au cas 2D et appliquée au traitement d'image et plus particulièrement à l'analyse des textures. Deux équipes sont connues pour avoir travaillé sur le sujet : Francos et al. [38, 70] et Liu et Picard [81].

L'image de texture est considérée comme étant la réalisation d'un champ aléatoire. Francos suppose de plus que ce champ est homogène et régulier. Dans un premier temps, ce champ est décomposé en un champ purement non déterministe et un champ déterministe. La composante déterministe est, par la suite, redécomposée en deux composantes orthogonales : la composante évanescence et la composante harmonique. Dans le domaine spectral, la composante purement non déterministe est absolument continue tandis que la composante déterministe est singulière par rapport à la mesure de Lebesgue.

La composante non déterministe est modélisée par un modèle AR. La composante harmonique est approchée par un champ aléatoire qui dans le domaine spectral est composé de fonctions de Dirac 2D. Quant à la composante évanescence, elle est approchée par un champ dont la représentation spectrale est une somme finie de fonctions de Dirac 1D supportées par des lignes de pente rationnelle. Francos et al. proposent de décomposer l'image en chacune de ces composantes par seuillage dans le domaine spectral et de les modéliser séparément. Selon les auteurs la supériorité de leur modèle sur ceux préexistants (modèles AR, MRF) réside dans le fait qu'aucune hypothèse n'est faite sur la distribution spectrale du champ aléatoire (comme par exemple qu'elle soit absolument continue).

Liu et Picard ont proposé un nouveau modèle basé sur un test d'harmonie. Ce test permet de caractériser l'image comme étant plus ou moins structurée et donc de calculer suivant les résultats obtenus les paramètres significatifs (coefficients du modèle AR et/ou les fréquences et les amplitudes des pics dans le domaine spectral). Ce modèle appliqué à la classification et à la segmentation d'images est intéressant dans la mesure où il permet de traiter des textures localement inhomogènes et est insensible à certaines transformations telles que les rotations.

### c. Modèles markoviens

L'ouvrage de Li [79] permet d'avoir un panorama assez vaste de l'utilisation des champs de Markov en vision par ordinateur, notamment dans le domaine de l'analyse et la synthèse de textures (pages 52-59). Le modèle le plus souvent utilisé est le champ de Markov gaussien ou GMRF (*Gaussian Markov Random Field*) [19, 21]. Dans [69], les auteurs définissent un modèle hiérarchique basé sur le modèle GMRF.

Dans le cas du GMRF classique, le signal observé au site  $s$  est égal à la somme, pondérée par les paramètres du modèle, des niveaux de gris des voisins de  $s$ , à laquelle on ajoute un bruit gaussien de moyenne nulle. Ce bruit, dont la variance constitue également un des paramètres du modèle, est très souvent estimé par la méthode des moindres carrés.

Dans [24], Cross et Jain utilisent le modèle binomial afin de synthétiser et caractériser des textures. La probabilité de chaque pixel de la texture conditionnellement à ces voisins est binomiale. Le paramètre de la loi dépend des voisins et du nombre de niveaux de gris.

Dans [29], les auteurs définissent deux modèles markoviens gaussiens l'un isotrope, l'autre anisotrope. Ils proposent deux méthodes pour estimer les paramètres caractéristiques de leurs modèles. L'une des méthodes est fondée sur la construction de matrices de probabilités conditionnelles. L'autre méthode est fondée sur la théorie de la renormalisation [44].

D'autres chercheurs se sont appliqués à définir de nouveaux modèles. Gimel'farb [45] propose une approche originale. Il estime à la fois les potentiels de son modèle et la structure, i.e. il n'a aucun a priori sur la forme de ses potentiels, ni sur la structure des voisinages. Il se limite aux interactions de deux pixels. Tous ces paramètres sont estimés à partir des histogrammes des différences des niveaux de gris. Son modèle suppose une structure invariante spatialement décrite par des familles de cliques invariantes par translation. Les potentiels sont approchés puis raffinés par une méthode appelée recuit simulé contrôlé qui s'inspire de la méthode du gradient stochastique proposée par Younes dans [131].

### d. Modèle fractal et multifractal

Batty et Longley [10] proposent d'utiliser la géométrie fractale afin d'étudier la

forme des villes. Selon eux, si la forme de la ville au premier coup d’œil apparaît irrégulière, la géométrie fractale est un outil qui permet de mettre en avant un certain “ordre” caractéristique de sa complexité et de sa diversité.

Dans les premières méthodes fractales, l’image fut considérée comme une surface dans un espace à trois dimensions, la troisième dimension étant le niveau de gris de l’image. Plusieurs équipes ont ainsi cherché à caractériser les textures à l’aide du calcul de la dimension fractale [97, 98], extension au cas 2D des études menées par Mandelbrot.

Dans [78], Levy-Véhel et al. mettent en doute le bien fondé de l’approche qui consiste à considérer une image de niveaux de gris comme une surface dans un espace 3D. Ils proposent dans le cadre de la théorie multifractale de décrire l’image par un triplet  $(\alpha, g(\alpha), l(\alpha))$  où  $\alpha$  est l’exposant de Hölder,  $g(\alpha)$  la dimension fractale de  $E_\alpha$  (ensemble des points ayant le même exposant), et  $l(\alpha)$  représente la lacunarité (mesure l’homogénéité d’une distribution fractale).

### 1.2.3.5 Méthodes mixtes

Certains chercheurs ont mis en évidence les liens qui existent entre les différentes méthodes décrites précédemment. Dans [99], Picard et al. ont décrit les liens théoriques qui existent entre les matrices de cooccurrence, la morphologie mathématique, et les modèles gibbsiens. Ce travail est présenté de manière plus détaillée dans la suite du manuscrit.

Prêteux et Descombes [100] utilisent le concept de processus hiérarchique en combinant processus booléens (simulant la macro-texture) et modèle markovien (modélisant la micro-texture). Ils permettent ainsi des ruptures de modèle ce qui autorise la représentation de textures très variées. Enfin, une méthode qui nous a paru intéressante est celle proposée par Zhu et al. [133] en 1996 qui allie filtrage et champs aléatoires.

### 1.2.3.6 Conclusion

Les méthodes d’analyse de texture sont riches et variées. Cependant, deux types de méthodes semblent revenir de manière récurrente dans la littérature sur le sujet : celles fondées sur le filtrage de l’image et plus particulièrement celles utilisant les filtres de Gabor et les ondelettes; celles fondées sur une modélisation de l’image et plus particulièrement sur une modélisation markovienne. Ajoutons, pour être complet, que les paramètres de texture estimés à partir des matrices de cooccurrence

sont souvent utilisés, même si cette approche est à notre sens peu souple comme nous le verrons dans ce qui suit. Elle n'en demeure pas moins une bonne méthode de référence, connue de tous, dans le cadre d'une étude comparative.



## DANS CE CHAPITRE:

2.1	Introduction . . . . .	26
2.2	Discrimination bâti/non bâti: état de l'art et applications . . . . .	26
2.3	Modèles isotropes . . . . .	29
2.3.1	Modèle markovien gaussien . . . . .	29
2.3.1.1	Définition . . . . .	29
2.3.1.2	Estimation des paramètres : "queues de comètes" . . . . .	30
2.3.2	Matrices <code>aura</code> . . . . .	31
2.3.3	Pourquoi prendre en compte l'anisotropie ? . . . . .	37
2.4	Modèles fondés sur des chaînes . . . . .	39
2.4.1	Introduction . . . . .	39
2.4.2	Les modèles . . . . .	40
2.4.3	Normalisation des paramètres . . . . .	42
2.4.4	Résultats . . . . .	50
2.4.5	Comparaison avec deux méthodes classiques . . . . .	53
2.4.5.1	Filtres de Gabor . . . . .	53
2.4.5.2	Matrices de cooccurrence . . . . .	64
2.4.6	Conclusion . . . . .	69

# 2.

## Masque Urbain: Analyse de texture

Comme nous l'avons souligné dans l'introduction la caractérisation des thèmes bâti/non bâti à partir d'images satellitales est nécessaire à de multiples applications. Ce type d'étude est utilisé pour la gestion des territoires : dans les pays émergents pour étudier la pression urbaine sur d'autres thèmes (forêt, agriculture...), dans les pays industrialisés pour gérer, par exemple, les surfaces industrialisées.

L'extraction du bâti nécessite, dans un premier temps, une analyse de la texture de l'image. Cette étape permet d'extraire des paramètres discriminant les thèmes bâti/non bâti. Nous proposons donc, dans ce chapitre, une méthode de caractérisation du thème bâti (par rapport au thème non bâti) à partir d'une analyse de la texture par modélisation markovienne.

## 2.1 Introduction

Dans ce qui suit nous rappelons les différentes approches qui existent pour extraire les thèmes bâti et non bâti dans des images satellitales.

Par la suite, nous détaillons deux méthodes utilisées pour analyser de manière isotrope la texture urbaine. La première est fondée sur un modèle markovien, la seconde sur les `matrices_aura`, dérivées des matrices de cooccurrence.

Après avoir montré les limites des méthodes précédemment décrites et la nécessité d'une analyse anisotrope, nous présentons nos travaux : modèle de texture, comparaison avec des méthodes classiques, algorithmes de classification, de régularisation.

Des résultats sur des images SPOT3, des simulations SPOT5 à diverses résolutions, et des images ERS1 sont présentés.

## 2.2 Discrimination bâti/non bâti: état de l'art et applications

La partition automatique d'image satellitale contenant des zones urbaines est particulièrement difficile à réaliser. La plupart du temps, l'information apportée par le seul niveau de gris est largement insuffisante<sup>1</sup>. Parmi les diverses approches utilisées, la plupart font appel à des paramètres de texture. Dans la majorité des nouveaux travaux sur l'extraction de l'urbain par une analyse de la texture les auteurs ne proposent pas de nouveaux paramètres. La nouveauté provient plutôt des algorithmes de segmentation ou d'estimation des paramètres, sachant que ces derniers sont classiquement les mêmes (issus des matrices de cooccurrence, indice de végétation ...).

Ainsi, Gouinaud dans sa thèse [47, 48] propose un comparatif des divers paramètres de texture utilisés pour la partition d'images satellitales optiques. Il montre bien qu'il n'existe pas de paramètre "idéal" mais plutôt un ensemble de paramètres "utiles". Certains d'entre eux étant aptes à distinguer relativement correctement l'urbain du non urbain (indice de température [25]) mais échouant lors de la distinction des différents types d'urbain, d'autres ayant des capacités inverses (indice de taille), d'autres encore échouant dans tous les cas (indice de brillance). Cette étude a orienté notre travail dans la mesure où nous avons choisi

---

<sup>1</sup>"*la couleur* ou l'aspect de [l'image satellitale], le ton, ne peuvent être pris en considération, seuls. La réflectance d'un objet ne permet pas en général d'en déterminer sa nature." [124] page 91.

de procéder par étape : extraction urbain/non urbain, analyse intra-urbaine.

Parmi les travaux effectués dans ce domaine, on citera ceux réalisés par Houzelle [59] qui fusionne l'information obtenue à l'aide de deux capteurs (SPOT et RSO). En effet, les agglomérations ont la particularité d'apparaître de manière brillante sur les images radar. Par un simple seuillage, il peut déterminer des masques d'apprentissage pour le calcul des paramètres des zones urbaines. Puis, par des considérations de distance par rapport aux centres de ces zones, il détermine des masques d'apprentissage pour les paramètres des autres zones. La partition est réalisée à l'aide d'un simple classifieur gaussien multivarié classique. Il utilise quatre attributs de texture (deux liés à l'histogramme local, une mesure locale du laplacien et une mesure d'orientation du gradient).

Haralick [55] a proposé des paramètres calculés à partir de la matrice de cooccurrence (entropie, énergie, corrélation  $\dots$ ). Il les a utilisés pour segmenter des images satellitales LANDSAT contenant une ou plusieurs classes d'urbain. Cependant, Gouinaud [47] ainsi que Baraldi et Parmiggiani [8] soulignent l'inadéquation de ces paramètres avec l'analyse du milieu urbain sur des images SPOT. Ceci est dû au fait que les images SPOT (notamment les images du futur capteur SPOT5) ont une bien meilleure résolution que les images LANDSAT et donc, qu'à cette résolution, la texture urbaine est d'échelle variable.

D'autres approches sont basées sur la modélisation de l'image. Les paramètres du modèle lorsqu'ils sont estimés sur la zone urbaine constituent alors les paramètres de texture [25]. Parmi les modélisations possibles de l'image, nous citerons l'approche markovienne qui est celle que nous avons utilisée dans nos travaux.

Hormis l'analyse de la texture, on rencontre également des méthodes fondées sur une analyse par morphologie mathématique. Ainsi, Serendero [108] utilise la transformation du chapeau haut de forme (image originale soustraite à son ouverture) afin d'extraire des amas de pixels traduisant la présence d'une ville. Pour colmater les zones urbaines elle applique ensuite un filtre alterné séquentiel (suite d'ouvertures et de fermetures de taille croissante). Son but est d'obtenir une localisation grossière de la ville et non pas une extraction précise. D'autres [10] utilisent la géométrie fractale pour étudier la forme urbaine.

Le CNES a, quant à lui, collaboré avec plusieurs laboratoires de recherche en vue de développer des méthodes d'extraction des zones urbaines<sup>2</sup>. Les approches proposées font appel aux opérateurs de contours (collaboration avec l'IRIT), à la morphologie

---

<sup>2</sup><http://www-projet.cst.cnes.fr:8090/agglomeration/menu.html>

mathématique [108] et aux réseaux neuromimétiques [125] (collaborations avec l'INRIA Sophia-Antipolis).

Enfin nous citerons également les travaux de Winter [129] qui a développé une méthode d'analyse multi-échelle fondée sur l'hypothèse qu'il existe pour chaque objet une échelle optimale de détection. La méthode FAME (Focalisation d'Attention Multi-échelle Entropique) utilise des mesures entropiques sur des pyramides d'ondelettes. Elle est appliquée à l'analyse du paysage urbain au niveau du bâtiment et également à la détection du contour urbain.

La connaissance du masque urbain est nécessaire à des applications qui s'orientent essentiellement vers la gestion du territoire. Ainsi dans [4] les principales applications sont des révisions de schémas directeurs<sup>3</sup>, des suivis d'évolution du contour urbain. L'auteur donne l'exemple de la cartographie de l'urbanisation dans le district de Tunis entre 1975 et 1991. La révision du schéma directeur de Clermont-Ferrand a également nécessité l'extraction d'un masque urbain.

Des études sont également menées dans certaines régions pour étudier la pression du milieu urbain sur d'autres milieux que sont la forêt, les champs, le littoral. C'est le cas de l'étude menée sur la ville de Tanger par SCOT Conseil en collaboration avec la faculté des sciences et techniques de Tanger<sup>4</sup>. De même dans [76], Letourneau et al. utilisent l'extraction d'un masque urbain sur des images ERS dans le but de suivre l'extension de la ville de Macapá (Brésil).

Soulignons, par ailleurs, que plusieurs organismes de recherche (INRIA, CNRS, BRGM etc.) vont collaborer sur un projet d'"utilisation des différents capteurs satellitaires pour la mise à jour des indicateurs liés aux problèmes de ressources hydrogéologiques et minières en Afrique". Le but essentiel est de développer des méthodes automatiques ou semi-automatique pour mettre à jour des cartes géographiques à partir d'images satellitaires. La zone d'étude se situe au nord-est de la Guinée-Conakry (frontière Guinée-Mali). La première étape consistera à valider les méthodes développées sur des zones favorables, i.e. sur lesquelles on peut acquérir des données terrain. Les intérêts de l'automatisation du processus de mise à jour sont multiples. Elle permettra notamment d'étudier des zones inaccessibles, éloignées des métropoles et peu documentées, ainsi que des zones vastes. Par ailleurs, l'évolution rapide de certains thèmes, e.g. la localisation des villes, nécessite un traitement automatisé des données. Un des buts ultimes est

---

<sup>3</sup>"déterminent les grandes orientations de l'évolution de l'urbanisme pour les territoires auxquels ils s'appliquent". Le Petit Larousse Illustré.

<sup>4</sup><http://www.scot-sa.com/tanger/>

de déterminer les besoins potentiels en eau à partir de critères indirects comme la localisation et la taille des villages.

L'extraction d'un masque urbain peut également être d'un grand intérêt dans le domaine de l'analyse adaptative d'images. En effet, Benharrosh [12] souligne la nécessité d'extraire des thèmes cartographiques précis dans le but soit de générer des *quick-looks* adaptatifs, soit de compresser des images de manière sélective. En effet, il n'est souvent pas nécessaire de disposer d'une haute résolution sur toute l'image. En revanche, certaines régions, qualifiées de zones d'intérêt par l'utilisateur, comme par exemple un réseau linéique ou encore une zone urbaine, doivent être préservées. Le problème de la détection et de l'extraction de ces zones n'est pas abordé par l'auteur.

Notons qu'en ce qui nous concerne, l'extraction du masque urbain est un préalable à la caractérisation intra-urbaine.

L'analyse du paysage urbain à partir d'images satellitales de manière automatique demeure pour l'instant peu appliquée. L'analyse par photo-interprétation assistée par ordinateur reste prépondérante.

## 2.3 Modèles isotropes

Dans [47], Gouinaud fait une analyse comparative très intéressante de nombreuses méthodes utilisées pour caractériser, puis pour extraire un masque urbain. Parmi celle-ci, le modèle markovien gaussien isotrope proposé par Descombes dans [25, 29] donne des résultats satisfaisants comparés aux autres approches. C'est pourquoi, dans un premier temps nous présentons le modèle en question.

### 2.3.1 Modèle markovien gaussien

#### 2.3.1.1 Définition

Dans [29] les auteurs proposent un nouveau paramètre de texture  $\beta$ . Ils définissent  $\beta$  en modélisant l'image par un modèle markovien gaussien 4 connexe (cf. équation 2.1). Dans tout ce qui suit, nous considérons des conditions aux bords toroïdales.

$$P(X_s|X_r, r \in V(s)) = \frac{1}{Z_{V(s)}} \exp \left( -\beta \left( \sum_{r \in V(s)} (X_s - X_r)^2 + \lambda(X_s - \mu)^2 \right) \right) \quad (2.1)$$

En réécrivant l'équation (2.1), on montre [25] que la distribution conditionnelle  $P(X_s|X_r, r \in V(s))$  dépend uniquement de la moyenne  $m_s$  des 4 voisins de  $s$  (cf. équations (2.2) et (2.3)). Les estimateurs sont alors plus robustes car ils ne dépendent plus que d'une seule variable conditionnante  $m_s$  (cf. équation (2.3)).

$$P(X_s|X_r, r \in V(s)) = \frac{1}{Z_{V(s)}} \exp \left( -\beta(4 + \lambda) \left( X_s - \frac{4m_s + \lambda\mu}{4 + \lambda} \right)^2 \right) \quad (2.2)$$

$$= P(X_s|m_s) \quad (2.3)$$

$\lambda$  et  $\beta$  sont les paramètres du modèle.

$\mu$  est la moyenne locale.

$V(s)$  est le voisinage de  $s$ .

$Z_{V(s)}$  est la constante de normalisation (fonction de partition locale).

Nous nous plaçons dans un cadre stationnaire, i.e. la moyenne locale  $\mu$  est indépendante de  $s$ .

La distribution conditionnelle locale  $P(X_s|X_r, r \in V(s))$  est une loi normale définie comme suit :

$$P(X_s|X_r, r \in V(s)) \equiv \mathcal{N} \left( \frac{4m_s + \mu\lambda}{4 + \lambda}, \frac{1}{2\beta(4 + \lambda)} \right) \quad (2.4)$$

La variance conditionnelle locale  $\sigma_{P(X_s|m_s)}^2 = \frac{1}{2\beta(4 + \lambda)}$  est estimée par la méthode des "queues de comètes" décrite dans le paragraphe suivant.

### 2.3.1.2 Estimation des paramètres : "queues de comètes"

Nous estimons les paramètres de texture, i.e. les variances conditionnelles locales, en utilisant la méthode des "queues de comètes" fondée sur l'estimation des matrices des probabilités conditionnelles. Cette méthode a été proposée pour le

première fois par Descombes dans [25], puis détaillée par la suite dans [28, 29]. Le nom “queues de comètes” provient de la forme des matrices qui rappelle celle des comètes.

La méthode consiste à estimer la matrice des probabilités conditionnelles dans une fenêtre glissante. Le choix de la taille de la fenêtre est le résultat d’un compromis. Plus la fenêtre est large, plus les statistiques sont robustes. D’un autre côté, plus la fenêtre est étroite, moins il y a de risques de se trouver en présence d’un mélange de plusieurs textures. Ensuite, pour chaque position de la fenêtre, on estime la variance conditionnelle locale à partir de la matrice des probabilités conditionnelles. Ce paramètre est assigné au pixel situé au centre de la fenêtre.

En pratique, pour chaque site  $s$  à l’intérieur de la fenêtre on incrémente de 1 l’élément de la matrice dont les coordonnées correspondent à  $X_s$  et à la moyenne  $m_s$ . Chaque ligne de la matrice représente une distribution conditionnelle  $P(X_s|m_s)$  qui est une loi normale (voir équation 2.4) dont la variance  $\sigma_{P(X_s|m_s)}^2$  ne dépend pas de  $m_s$ .

On estime donc  $\sigma_{P(X_s|m_s)}^2$ , par son estimateur empirique, à partir de la distribution  $P(X_s|m_s)$  représentée par le plus grand nombre d’échantillons afin d’obtenir un estimateur robuste et des bords francs. En effet, dans les cas où la fenêtre d’estimation recouvre deux textures différentes, de moyennes nettement différentes, nous ne tenons compte que de la texture la plus représentée [28].

Cette méthode nous permet d’estimer la variance conditionnelle locale.

Comme nous le verrons par la suite (2.4.3) nous aurons également besoin d’estimer le paramètre  $\lambda$ .

Nous remarquons que la moyenne de la loi  $P(X_s|m_s)$  est une fonction linéaire de  $m_s$  (cf. équation 2.4), i.e.  $\langle P(X_s|m_s) \rangle = a * m_s + b$ , où  $a = \frac{4}{4+\lambda}$  et  $b = \frac{\lambda\mu}{4+\lambda}$ .

Nous estimons donc  $\lambda$  en minimisant l’erreur quadratique  $\sum_{(x=m_s, y=\langle \cdot \rangle)} (y - ax - b)^2$

par la méthode des moindres carrés. Cette minimisation est effectuée par rapport à  $a$  et  $b$ . Nous déduisons  $\lambda$  de  $a$  et  $b$ .

### 2.3.2 Matrices aura

Afin de comparer l’analyse décrite précédemment à des méthodes plus classiquement utilisées, nous présentons les **ensembles aura**, **mesures aura**, et les **matrices de miscibilité** proposées par Picard et Elfadel dans [35, 99]. Cette approche basée sur la théorie des ensembles a le grand mérite de lier les matrices de cooccurrence, les champs de Gibbs, et la morphologie mathématique. De plus, alors que les ma-

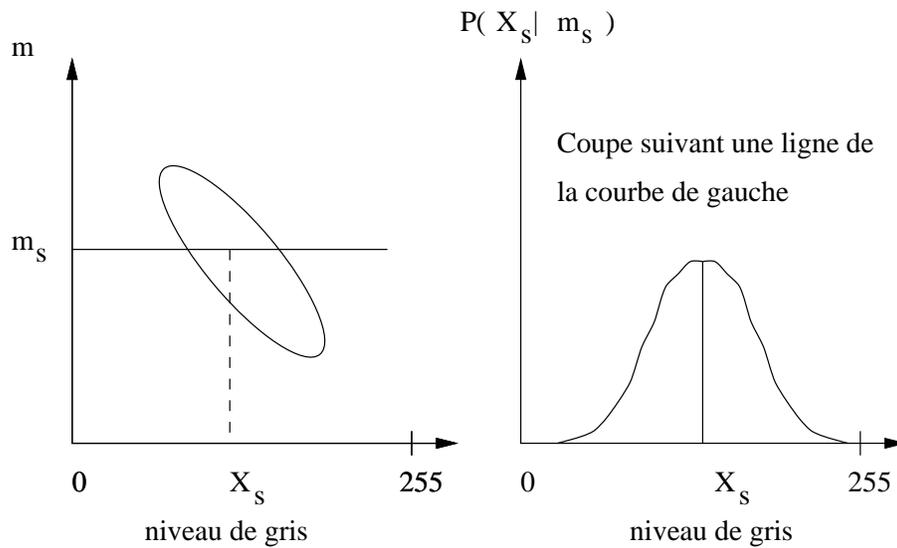


Figure 2.1: Queue de comète.

trices de cooccurrence introduisent (par le vecteur de déplacement) une notion de direction, les matrices de miscibilité peuvent être définies sur un voisinage isotrope. Elles constituent donc un cadre comparatif très intéressant.

### a. Définitions

Dans la suite, nous noterons  $S$  un réseau fini de sites et  $V = \{V(s) | s \in S\}$  le système de voisinage associé ( $V(s)$  est le voisinage du site  $s$ ).

**Définition 2.3.1** Soient  $A$  et  $B$  deux sous ensembles de  $S$  (Figure 2.2). L'aura de  $A$  par rapport à  $B$  pour le système de voisinage  $V$  est un sous ensemble de  $S$  défini comme suit :

$$\bigcup_{s \in A} (V(s) \cap B)$$

L'aura de  $A$  par rapport à  $B$  sera noté  $\mathcal{O}_B(A, V)$ .



Dans la suite, on notera  $m_a(g, g')$  au lieu de  $m_a(S_g, S_{g'})$ .

**Définition 2.3.4** *Les matrices de miscibilité  $\mathbf{M}_m$  sont définies comme les matrices  $\mathbf{M}_a$  normalisées.*

$$\mathbf{M}_m = [\mathbf{m}_m(\mathbf{g}, \mathbf{g}')] = \frac{|\mathbf{G}|}{|\mathbf{V}(s)||\mathbf{S}|} \mathbf{M}_a$$

Dans [35], les auteurs interprètent  $m_m(g, g')$  comme une probabilité.

$m_m(g, g')$  est la probabilité qu'un site de niveau de gris  $g'$  ait dans son voisinage un site de niveau de gris  $g$  :  $m_m(g, g') = P(\exists r \in V(s), X_r = g | \exists s \in S, X_s = g')$ .

## b. Quels liens avec les autres approches ?

L'intérêt essentiel de ce concept est de permettre de montrer les relations qui existent entre différentes approches (Champs de Gibbs, *Gibbs Random Fields*; matrice de cooccurrence; morphologie mathématique) très souvent considérées séparément pour la classification d'images de texture.

L'élément de base des transformations par morphologie mathématique [106] est l'élément structurant que nous notons  $\mathcal{V}$ .

$\mathcal{V}_s$  désigne l'élément structurant centré sur le pixel  $s$ . L'analyse par morphologie mathématique consiste à déplacer  $\mathcal{V}$  sur la trame de l'image et à étudier les interactions entre les objets de l'image et  $\mathcal{V}$ . Pour plus de détails, le lecteur peut se reporter au chapitre 4.

**Définition 2.3.5** *La dilatation de  $A$  par  $\mathcal{V}$ , notée  $A \oplus \check{\mathcal{V}}$  est définie de la manière suivante :*

$$A \oplus \check{\mathcal{V}} = \{s | \mathcal{V}_s \cap A \neq \emptyset\} = \bigcup_{t \in \check{\mathcal{V}}} A_t$$

$A_t$  est le translaté de l'ensemble  $A$  par le vecteur  $t$ .

$\check{\mathcal{V}}$  est le symétrique de  $\mathcal{V}$  par rapport à l'origine. Dans la suite, nous considérons des éléments structurants symétriques par rapport à l'origine donc tels que  $\check{\mathcal{V}} = \mathcal{V}$ .

Picard et Elfadel expriment l'aura de  $A$  par rapport à  $B$  en terme de dilatation :

$$\begin{aligned}\mathcal{O}_B(A, V) &= \bigcup_{s \in A} (V(s) \cap B) \\ &= \left( \bigcup_{s \in A} V(s) \right) \cap B \\ &= (\mathcal{V} \oplus A) \cap B\end{aligned}$$

$\mathcal{V}$  est l'élément structurant.  $V(s)$  est donc une version translatée de  $\mathcal{V}$ .

$\mathcal{V} \oplus A$  est la dilatation de  $\mathcal{V}$  par  $A$ .

Les **matrices aura** sont des sommes de matrices de cooccurrence obtenues pour des vecteurs de déplacement  $\mathbf{d}_k$  contenus dans l'élément structurant.

Ainsi dans le cas où  $V(s) = \{s + \mathbf{d}\}$  alors  $c_{gg'}(\mathbf{d}) = \mathbf{m}_a(\mathbf{g}, \mathbf{g}')$ .

Dans le cas général :

$$m_a(g, g') = \sum_{k=1}^K c_{gg'}(\mathbf{d}_k) \quad (2.5)$$

Enfin, les auteurs reformulent les fonctions d'énergie issues de la modélisation par les GRF (*Gibbs Random Fields*) comme combinaison linéaire des **mesures aura**.

En faisant l'hypothèse d'un champ homogène et isotrope<sup>5</sup>, la fonction non linéaire  $\sum_{s \in S} \sum_{r \in V(s)} \mathbf{V}(X_s, X_r)$ , où  $\mathbf{V}(X_s, X_r)$  est le potentiel d'interaction entre

$X_s$  et  $X_r$ , est réécrite en utilisant les **mesures aura** de la manière suivante :  $\sum_{g, g' \in G} \mathbf{V}(g, g') m_a(g, g')$ . Cette reformulation permet notamment de transformer

un problème d'optimisation d'un GRF dans le domaine des niveaux de gris en un problème d'optimisation dans le domaine de miscibilité. Dans [35], les auteurs donnent l'exemple de l'échantillonnage du modèle de Potts avec la contrainte d'un histogramme uniforme (cette dernière hypothèse est assez discutable). Le problème d'optimisation formulé dans le domaine de miscibilité est :

$$\min_{x \in \Omega} \left( \sum_{g \in G} m_a(g, g') - 2 \sum_{g < g'} m_a(g, g') \right),$$

sous la contrainte  $|S_g| = \frac{|S|}{|G|}, \forall g \in G$ .

---

<sup>5</sup>Notons que dans le cas d'un champ anisotrope, Picard et Elfadel décomposent le voisinage anisotrope en un ensemble de sous voisinages, de tel façon que le champ se comporte de manière isotrope dans chaque sous voisinage

Dans [35], les auteurs donnent également l'exemple de la reformulation du problème d'optimisation dans le cas du modèle autobinominal [14, 24].

Ce concept nous paraît très intéressant car il permet de mettre en exergue les liens entre les différentes approches considérées séparément la plupart du temps. Dans le paragraphe suivant, nous avons utilisé ce concept pour comparer les résultats obtenus avec le modèle markovien gaussien isotrope 2D et les `matrices de miscibilité`.

### c. Résultats expérimentaux

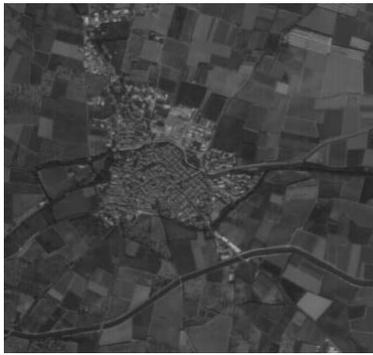
Des `matrices de miscibilité`, on peut extraire des paramètres comme des matrices de cooccurrence. Pour le problème considéré dans ce manuscrit, nous avons utilisé l'entropie et l'énergie car ces deux paramètres étaient les plus discriminants. Soulignons que dans les tests qui suivent le choix de la taille des fenêtres est empirique. Cette taille doit être suffisamment grande pour avoir des statistiques robustes et suffisamment petite pour que la fenêtre contienne une texture homogène. Nous avons testé plusieurs tailles de fenêtre et nous avons conservé celle qui offrait le meilleur compromis.

Pour l'image test présentée figure 2.3, nous estimons le paramètre dérivé du modèle markovien gaussien 2D sur une fenêtre de taille 17x17.

Nous estimons également les matrices de miscibilité desquelles nous déduisons les deux images de paramètres (entropie et énergie) sur une fenêtre de taille 9x9. Enfin, nous avons également voulu comparer la variance conditionnelle locale avec la variance classique estimée sur la même fenêtre.

Le voisinage considéré est le voisinage 4 connexe. D'autre part, avant d'estimer les `matrices de miscibilité`, nous avons ramené l'image sur 32 niveaux de gris par une simple transformation linéaire. En effet, nous devons stocker des matrices dont le nombre d'éléments est égal au carré du nombre de niveaux de gris. Il faut donc diminuer ce nombre afin de réduire la taille des matrices à stocker. Notons que le fait d'utiliser davantage de niveaux de gris n'améliore pas les résultats. C'est également le cas lorsque l'on estime des matrices de cooccurrence.

L'intérêt de modéliser le signal correspondant à l'image est de pouvoir par la suite utiliser des méthodes pour estimer les paramètres du modèle. Dans notre cas la méthode des "queues de comète" nous permet d'obtenir des bords francs entre les classes. Cette différence apparaît nettement lorsque nous comparons les images



**Figure 2.3:** *Image initiale : Toreilles.*

de paramètre (Figures 2.4 (a), (b) et (c)). Sur les images des paramètres extraits des matrices de miscibilité, les bords des régions sont plus flous. Soulignons, par ailleurs, que même si la dynamique de l'image est ramenée à 32 niveaux de gris, nous devons construire les matrices de miscibilité avant d'estimer les paramètres (énergie, entropie). Cette étape est coûteuse en temps de calcul.

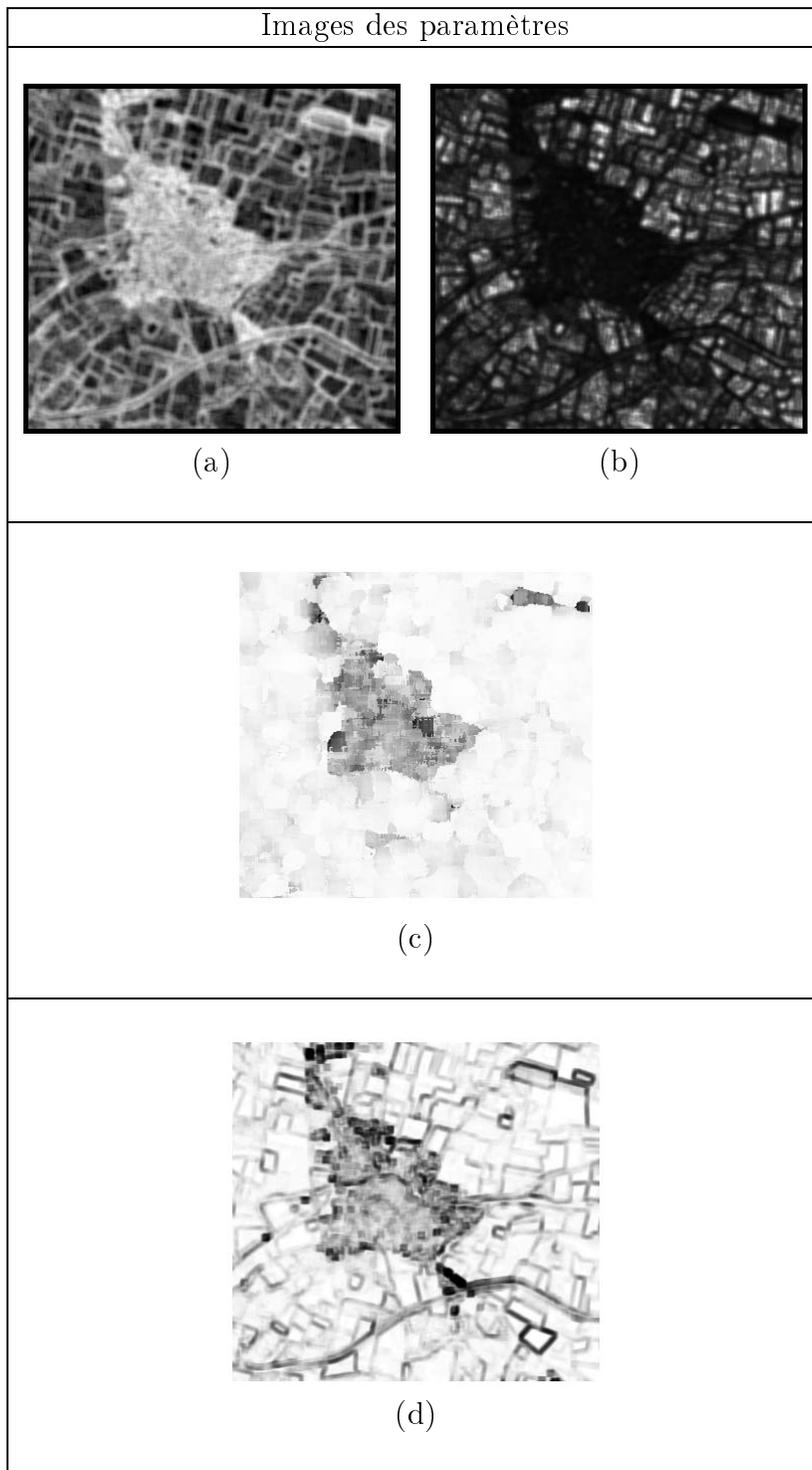
Dans le cas de l'estimation des variances conditionnelles locales, à aucun moment nous ne stockons la matrice des probabilités conditionnelles locales.

Enfin, il faut noter que la probabilité conditionnelle locale ne dépend que d'une variable conditionnante : la moyenne des niveaux de gris des voisins, ce qui rend les estimateurs de ces probabilités plus robustes.

Les résultats obtenus par le modèle markovien gaussien 4 connexe sont de meilleure qualité : paramètre plus discriminant, bords plus francs entre les classes, moins de données à stocker et notamment aucune matrice. Par ailleurs, il y a moins de fausses alarmes dues aux contours avec le modèle markovien.

### 2.3.3 Pourquoi prendre en compte l'anisotropie ?

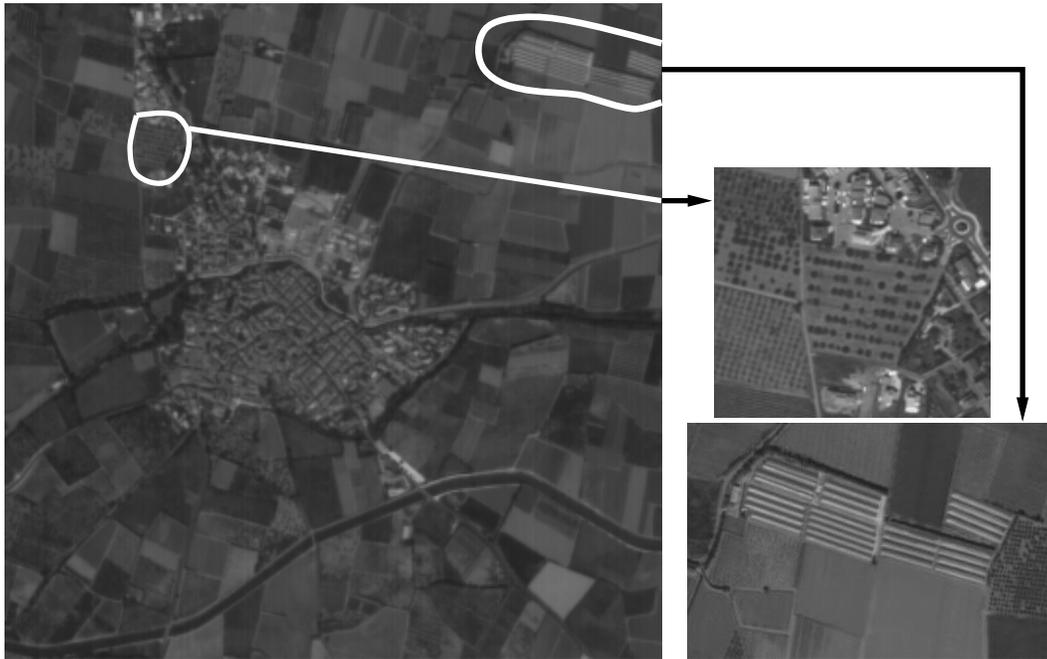
L'approche proposée est basée sur les résultats précédents et sur les observations suivantes : les images de zones urbaines contiennent des variations d'intensité des pixels quelque soit la direction, tandis que les images de champs sont constituées de régions d'intensité localement uniforme. Par ailleurs, certains objets tels que les serres, les champs ayant une structure linéaire (Figure 2.5) présentent des variations d'intensité dans des directions privilégiées. L'anisotropie de certaine zones de l'image est de plus en plus visible avec l'amélioration de la résolution des images satellitaires. Sur l'image de la figure 2.5, on voit nettement les cultures en lignes, de même que



**Figure 2.4:** Résultats comparatifs sur l'image de Toreilles : (a) entropie, (b) énergie, (c) variance conditionnelle locale, (d) variance.

la structure linéaire des serres.

Ainsi, sur l'image de l'estimée de la variance conditionnelle locale (Figure 2.4 (c)), qui correspond à une analyse isotrope de la texture, les objets ayant une structure linéaire (serres, certaines cultures) ont un paramètre de texture proche de celui des villes (zones sombres sur l'image 2.4 (c)). C'est pour cette raison que nous proposons d'analyser la texture par une approche anisotrope afin de pouvoir mieux caractériser les zones urbaines sur des images sur lesquelles la structure des objets apparaît plus nettement.



**Figure 2.5:** Zooms sur l'image de Toreilles : structures anisotropes (serres, cultures linéaires).

## 2.4 Modèles fondés sur des chaînes

### 2.4.1 Introduction

Dans [29] les auteurs décrivent un modèle markovien gaussien anisotrope et proposent des techniques pour estimer les paramètres associés. Le modèle ne tient compte que de deux directions (verticale et horizontale).

Dans ce cas, la propriété markovienne s'écrit de la manière suivante :

$$\begin{aligned}
 P(X_s | X_r, r \in V(s)) &= P(X_s | m_s^h, m_s^v) \\
 &= \frac{1}{Z_{V(s)}} \exp \left( -\beta(4 + \lambda) \left( X_s - \frac{2\alpha_h m_s^h + 2\alpha_v m_s^v + \lambda\mu}{4 + \lambda} \right)^2 \right)
 \end{aligned}
 \tag{2.6}$$

$m_s^h$  et  $m_s^v$  sont les moyennes des voisins de  $s$  dans les directions horizontale et verticale.

$\alpha_h$  et  $\alpha_v$  sont les paramètres du modèle. Dans le cas isotrope  $\alpha_h = \alpha_v = 1$ .

Si l'on introduit d'autres directions, le modèle devient plus complexe. Notamment l'estimation des paramètres par les méthodes décrites dans [29] devient problématique. En effet, avec la méthode des "queues de comète", les auteurs travaillent successivement sur l'ensemble des pixels dont la moyenne des voisins horizontaux (respectivement verticaux) est égale à la valeur la plus représentée dans l'image, ce qui réduit l'ensemble des statistiques. En utilisant la méthode de renormalisation par décimation, ce problème ne se pose pas. Cependant, cette méthode est difficilement généralisable à des topologies plus complexes que la 4 connexité (dans ces cas il faut négliger certaines cliques, ou utiliser des processus de décimation complexes).

C'est pourquoi nous avons opté pour la solution décrite dans la suite, qui consiste à modéliser le champ de luminance par des modèles différents (un par direction considérée) fondés sur des chaînes.

## 2.4.2 Les modèles

Nous avons tenu compte de l'anisotropie en considérant 8 modèles différents fondés sur des chaînes (un modèle pour chaque direction) [82, 83]. Au lieu de considérer les voisinages classiques isotropes (4 connexité, 8 connexité), nous définissons dans la direction  $d$ , le voisinage du pixel  $s$  comme étant l'ensemble constitué des deux voisins les plus proches de  $s$  dans cette direction (Figure 2.6). Chaque modèle, défini dans une direction donnée, peut être apparenté à un modèle joint fondé sur des chaînes ayant une structure markovienne d'ordre 1.

Nous considérons des champs gaussiens de sorte que la probabilité condition-

nelle pour le modèle défini dans la direction  $d$  est la suivante :

$$P(X_s|X_r, r \in V^d(s)) = \frac{1}{Z_{V^d(s)}} \exp \left( -\beta^{(d)} \left( \sum_{r \in V_s^d} (X_s - X_r)^2 + \lambda^{(d)} (X_s - \mu)^2 \right) \right) \quad (2.7)$$

$V^d(s)$  est le voisinage de  $s$  défini dans la direction  $d$ .

$Z_{V^d(s)}$  est la fonction de partition locale.

$\beta^{(d)}$ ,  $\lambda^{(d)}$ , et  $\mu$  sont les paramètres du modèle.

En réécrivant l'équation (2.7), on montre que  $P(X_s|X_r, r \in V^d(s))$  ne dépend que de la moyenne  $m_s^d$  des deux voisins de  $s$  dans la direction  $d$  (cf. équations (2.8) et (2.9)) :

$$P(X_s|X_r, r \in V^d(s)) = \frac{1}{Z_{V^d(s)}} \exp \left( -\beta^{(d)} (2 + \lambda^{(d)}) \left( X_s - \frac{2m_s^d + \lambda^{(d)}\mu}{2 + \lambda^{(d)}} \right)^2 \right) \quad (2.8)$$

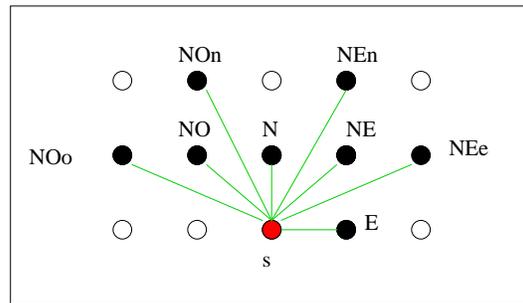
$$= P(X_s|m_s^d) \quad (2.9)$$

$P(X_s|X_r, r \in V^d(s))$  suit donc une loi normale définie par :

$$P(X_s|X_r, r \in V^d(s)) \equiv \mathcal{N} \left( \frac{2m_s^d + \mu\lambda^{(d)}}{2 + \lambda^{(d)}}, \frac{1}{2\beta^{(d)}(2 + \lambda^{(d)})} \right) \quad (2.10)$$

Pour chaque modèle nous estimons, les variances conditionnelles  $\sigma_d^2 = \frac{1}{2\beta^{(d)}(2 + \lambda^{(d)})}$  par la méthode des queues de comète [29].

Soulignons que notre objectif est d'extraire des paramètres de texture et non de synthétiser des textures. Il n'y a donc pas d'incohérence à considérer huit modèles différents. En fait, nous pouvons interpréter cette approche de la manière suivante : nous essayons d'ajuster les huit modèles de façon indépendante; chacun d'eux est supposé caractériser certaines propriétés du champ de luminance qui ne sont pas décrites par les autres modèles. De manière plus générale, les modèles les plus aptes à modéliser une texture en vue par exemple de la resynthétiser et ceux les plus aptes à discriminer des textures différentes en vue de classifier/segmenter ne sont pas nécessairement les mêmes. Plus concrètement un modèle de texture permettant de synthétiser une texture ne sera pas nécessairement le mieux adapté



**Figure 2.6:** Représentation des 8 directions prises en compte.

pour segmenter une image comportant plusieurs régions de textures homogènes.

Le réseau qui est constitué de la trame de l'image n'est pas isotrope comme le montre la figure 2.6. Si l'on considère la direction (E/O), les pixels voisins du pixel  $s$  sont situés à une distance 1 de  $s$ . En revanche, si l'on considère la direction (NOo/SEe), les pixels voisins de  $s$  sont situés à une distance  $\sqrt{5}$  du pixel  $s$ . Afin de corriger le biais introduit par l'anisotropie du réseau et de pouvoir comparer les paramètres estimés dans les huit directions, nous les normalisons.

### 2.4.3 Normalisation des paramètres

Le principe de renormalisation issu de la physique statistique [18, 60] consiste à réduire le nombre de degrés de liberté d'un modèle par changement d'échelle. La renormalisation a été utilisée notamment pour étudier les phénomènes se produisant au voisinage des transitions de phase [127]. Ces transitions sont caractérisées par des paramètres : les exposants critiques. Les états critiques d'un système sont caractérisés par une longueur de corrélation qui diverge et par conséquent qui est inchangée par changement d'échelle. En revanche, au voisinage d'un point critique la longueur de corrélation est finie et par conséquent devient plus petite par changement d'échelle (i.e. divisée par le pas de la transformation). Ceci signifie que le système s'éloigne de l'état critique. Ainsi, l'étude des points fixes instables de la transformation par renormalisation permet de déterminer les points critiques du système et aussi par exemple de calculer les exposants critiques.

Le principe de renormalisation a également été utilisé en traitement d'image. Dans [44], l'auteur utilise l'approche de groupe de renormalisation afin d'effectuer

une analyse multiéchelle de l'image. L'intérêt de cette technique réside, selon lui, dans le fait que les échelles d'analyse sont liées entre elle par une transformation non linéaire appelée *RG transformation* (*Renormalization Group transformation*). L'approche utilisée est appelée renormalisation par blocs. Elle consiste à remplacer un bloc de 4 pixels par son centre qui prend comme valeur le niveau de gris majoritaire parmi les 4. En cas d'égalité, le choix se fait en tirant suivant une loi uniforme parmi les niveaux de gris concurrents. D'autres solutions peuvent être envisagées, comme celle qui consiste à remplacer le bloc par son centre et à lui associer comme valeur la moyenne des 4 niveaux de gris. L'auteur applique cette méthode à la restauration d'image. Dans [90], Matsuba utilise la technique de groupe de renormalisation pour segmenter des images. Dans [29], les auteurs estiment les paramètres de leurs modèles en utilisant une méthode de renormalisation par décimation. Cette technique consiste à calculer les lois marginales sur un sous-réseau en intégrant la loi de probabilité associée au réseau initial sur les sites décimés.

La méthode de renormalisation par décimation va nous permettre de corriger l'anisotropie du réseau de pixels. Pour normaliser, nous choisissons un modèle défini sur le réseau avec un pas d'échantillonnage de référence. Puis, nous calculons les paramètres des différents modèles en fonction des paramètres du modèle défini sur ce réseau de référence.

Pour cela, nous intégrons le modèle de référence par rapport aux sites que l'on souhaite supprimer pour obtenir le réseau fourni par les données, ce qui revient à calculer la loi marginale sur ce dernier.

Notons  $\beta^{(d)}$  et  $\lambda^{(d)}$  les paramètres du modèle de référence.

Notons  $\beta_k^{(d)}$  et  $\lambda_k^{(d)}$  les paramètres du modèle décimé.

Dans ce qui suit, nous écrirons  $\beta_k$  et  $\lambda_k$  au lieu de  $\beta_k^{(d)}$  et  $\lambda_k^{(d)}$  pour simplifier les notations. Ainsi, sur la figure 2.7, les paramètres  $\beta_1$  et  $\lambda_1$  (l'indice est égal au nombre de variables sur lesquelles on intègre) sont obtenus à partir des paramètres  $\beta$  et  $\lambda$  en intégrant le modèle de référence par rapport à un pixel sur deux du réseau de référence.

Comme nous cherchons à calculer les paramètres de modèles dont les pixels sont distants de 1,  $\sqrt{2}$  ou  $\sqrt{5}$  par la méthode ci-dessus, il faut approcher ces nombres non forcément rationnels par des fractions rationnelles [107] afin que nous puissions intégrer sur un nombre entier de sites. D'autre part, il est nécessaire que ces fractions aient le même dénominateur afin de pouvoir définir le pas  $\epsilon$  du réseau de référence. Nous conservons donc les valeurs  $\frac{17}{12}$  et  $\frac{28}{12}$  pour approcher

respectivement  $\sqrt{2}$  et  $\sqrt{5}$ . Ces valeurs sont le meilleur compromis entre une bonne approximation et un nombre réduit de variables sur lesquelles intégrer. Dans ce cas,  $\epsilon = \frac{1}{12}$ .

$$\left| \frac{\sqrt{2} - \frac{17}{12}}{\sqrt{2}} \right| \approx 0.18\% < 5\%$$

$$\left| \frac{\sqrt{5} - \frac{28}{12}}{\sqrt{5}} \right| \approx 4.4\% < 5\%$$

Les racines carrées sont approchées par une suite de fractions rationnelles. Par ailleurs, nous devons approcher  $\sqrt{2}$ ,  $\sqrt{5}$  par des fractions rationnelles ayant le même dénominateur. Compte tenu de ces contraintes, l'approximation suivante pour  $\sqrt{2}$  est  $\frac{577}{408}$ . Ceci implique que nous intégrons le modèle de référence sur un très grand nombre de variables, en l'occurrence 576 (dans ce cas  $\epsilon = \frac{1}{408}$ ). C'est pourquoi nous choisissons  $\epsilon = \frac{1}{12}$  afin d'intégrer sur 16 et 27 variables.

$$\begin{aligned} 1 &= 12\epsilon \\ \sqrt{2} &\approx 17\epsilon \\ \sqrt{5} &\approx 28\epsilon \end{aligned}$$

Nous associons le couple de paramètres  $(\beta, \lambda)$  au réseau de référence de pas  $\epsilon$ . En intégrant le modèle de référence, nous obtenons la relation mathématique entre les paramètres estimés et les paramètres équivalents sur le réseau de référence. Nous estimons les paramètres sur l'image (suivant la direction considérée). Puis, nous recherchons les paramètres équivalents pour le réseau de pas  $\epsilon$ . Voici les résultats obtenus pour certaines de ces variables. Le détail des calculs pour le couple  $(\beta_1, \lambda_1)$  est proposé dans l'annexe. Dans ce cas, les calculs sont simples car les intégrales sont séparables.

Pour notre application, nous avons besoin de  $(\beta_k, \lambda_k)$ , pour  $k \in \{11, 16, 27\}$ . Nous

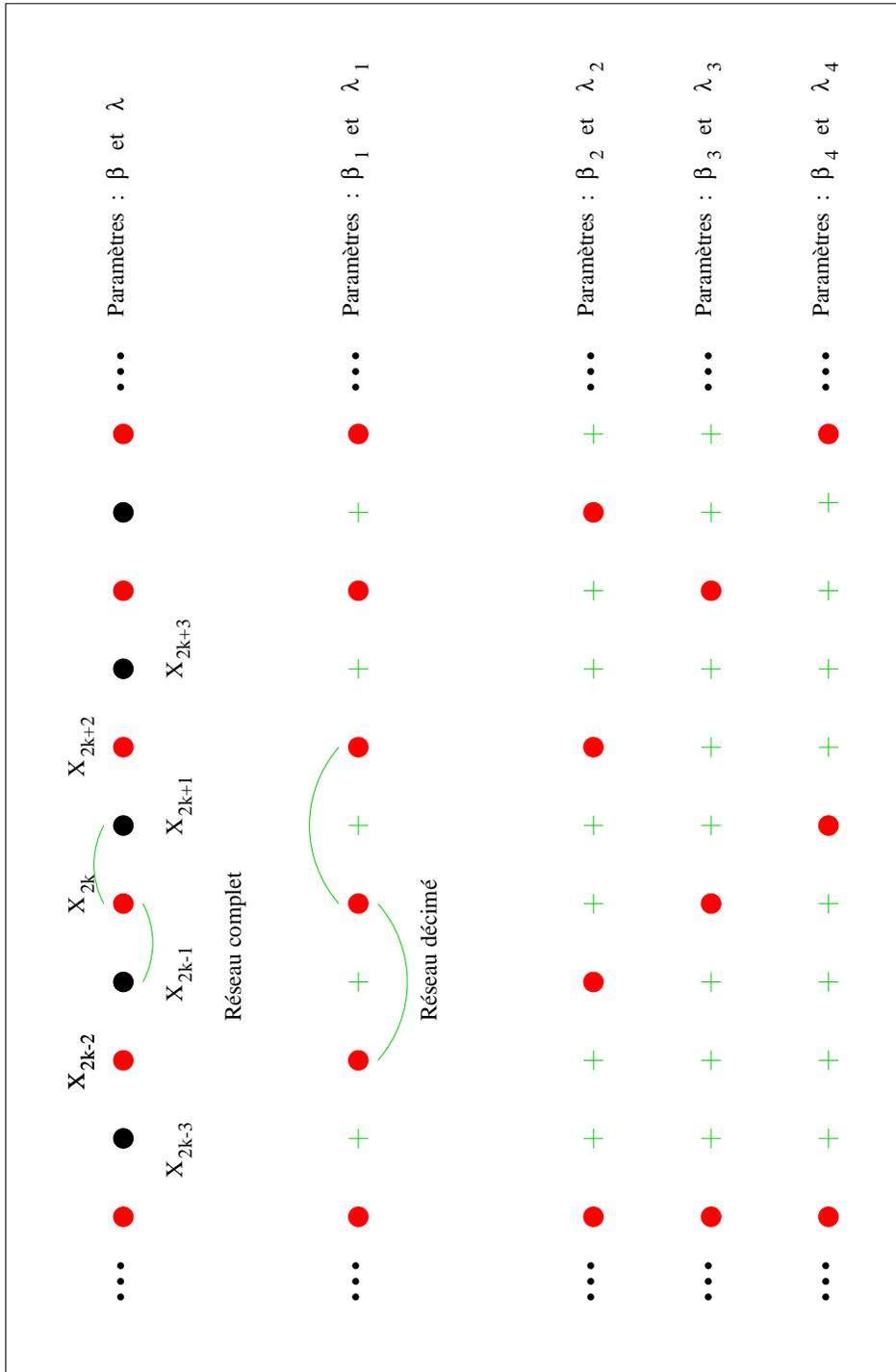


Figure 2.7: Représentation du réseau complet et des réseaux décimés.

utilisons le logiciel *Maple* pour les calculer.

La loi jointe pour le réseau de référence est la suivante :

$$P(x_1, x_2, \dots, x_N) \propto \exp \left( -\beta \left( \sum_{k=1}^N \lambda(x_k - \mu)^2 + \sum_{k=1}^{N-1} (x_k - x_{k+1})^2 \right) \right) \quad (2.11)$$

Nous intégrons cette loi par rapport à 11, 16 puis 27 variables. Les paramètres sont les suivants :

$$\lambda_{11} = \lambda(\lambda + 1)^2(\lambda + 2)^2(\lambda + 3)^2(\lambda + 4)(\lambda^2 + 4\lambda + 1)^2 \quad (2.12)$$

$$\lambda_{16} = \lambda(\lambda^8 + 17\lambda^7 + 119\lambda^6 + 442\lambda^5 + 935\lambda^4 + 1122\lambda^3 + 714\lambda^2 + 204\lambda + 17)^2 \quad (2.13)$$

$$\begin{aligned} \lambda_{27} &= \lambda(\lambda + 2)^2(\lambda + 4)(\lambda^3 + 7\lambda^2 + 14\lambda + 7)^2(\lambda^3 + 5\lambda^2 + 6\lambda + 1)^2 \\ &* (\lambda^6 + 12\lambda^5 + 53\lambda^4 + 104\lambda^3 + 86\lambda^2 + 24\lambda + 1)^2 \end{aligned} \quad (2.14)$$

$$\beta_{11} = \frac{\beta}{P_{11}(\lambda)}, \beta_{16} = \frac{\beta}{P_{16}(\lambda)} \text{ et } \beta_{27} = \frac{\beta}{P_{27}(\lambda)} \text{ où :}$$

$$P_{11}(\lambda) = (\lambda + 1)(\lambda + 2)(\lambda + 3)(\lambda^2 + 4\lambda + 1)(\lambda^2 + 4\lambda + 2)(\lambda^4 + 8\lambda^3 + 20\lambda^2 + 16\lambda + 1) \quad (2.15)$$

$$\begin{aligned} P_{16}(\lambda) &= (\lambda^8 + 17\lambda^7 + 119\lambda^6 + 442\lambda^5 + 935\lambda^4 + 1122\lambda^3 + 714\lambda^2 + 204\lambda + 17) \\ &* (\lambda^8 + 15\lambda^7 + 91\lambda^6 + 286\lambda^5 + 495\lambda^4 + 462\lambda^3 + 210\lambda^2 + 36\lambda + 1) \end{aligned} \quad (2.16)$$

$$\begin{aligned} P_{27}(\lambda) &= (\lambda + 2)(\lambda^2 + 4\lambda + 2)(\lambda^6 + 12\lambda^5 + 53\lambda^4 + 104\lambda^3 + 86\lambda^2 + 24\lambda + 1) \\ &* (\lambda^{12} + 24\lambda^{11} + 252\lambda^{10} + 1520\lambda^9 + 5813\lambda^8 + 14672\lambda^7 + 24648\lambda^6 + 27104\lambda^5 \\ &\quad + 18646\lambda^4 + 7344\lambda^3 + 1400\lambda^2 + 96\lambda + 1)(\lambda^3 + 7\lambda^2 + 14\lambda + 7)(\lambda^3 + 5\lambda^2 + 6\lambda + 1) \end{aligned} \quad (2.17)$$

Ainsi, si nous considérons les directions ( $d = \text{NOo}/\text{SEe}$ ,  $d = \text{NO n}/\text{SEs}$ ,  $d = \text{NE n}/\text{SOs}$ ,  $d = \text{NEe}/\text{SOo}$ ), les pixels voisins de  $X_s$  sont situés à une distance  $\sqrt{5}$  de celui-ci.

La variance conditionnelle estimée sur l'image dans ces directions est :

$$\begin{aligned}\sigma_d^2 &= \frac{1}{2\beta_{27}(d)(2 + \lambda_{27}(d))} \\ &= \frac{P_{27}(\lambda)}{2\beta(2 + \lambda_{27}(d))}\end{aligned}$$

La variance conditionnelle estimée sur l'image dans les directions ( $d = \text{NE}/\text{SO}$ ,  $d = \text{NO}/\text{SE}$ ) est :

$$\begin{aligned}\sigma_d^2 &= \frac{1}{2\beta_{16}(d)(2 + \lambda_{16}(d))} \\ &= \frac{P_{16}(\lambda)}{2\beta(2 + \lambda_{16}(d))}\end{aligned}$$

La variance conditionnelle estimée sur l'image dans les directions ( $d = \text{N}/\text{S}$ ,  $d = \text{E}/\text{O}$ ) est :

$$\begin{aligned}\sigma_d^2 &= \frac{1}{2\beta_{11}(d)(2 + \lambda_{11}(d))} \\ &= \frac{P_{11}(\lambda)}{2\beta(2 + \lambda_{11}(d))}\end{aligned}$$

En pratique, nous ne travaillons pas sur les couples  $(\beta, \lambda)$  mais sur les variances conditionnelles (qui sont des fonctions de  $\beta$  et  $\lambda$ ). De plus, nous ne chercherons pas les paramètres équivalents relativement au réseau de référence (correspondant à une distance de  $\epsilon$  entre 2 sites voisins) pour les trois distances ( $\sqrt{2}$ ,  $\sqrt{5}$ , 1). Mais nous allons chercher, pour les distances  $\sqrt{2}$  et  $\sqrt{5}$ , la variance conditionnelle équivalente relativement au réseau de pas 1. Ainsi, nous conservons les estimées des variances conditionnelles dans les directions N/S et E/O et nous normalisons les estimées dans toutes les autres directions. Le réseau de référence (correspondant à  $\epsilon$ ) n'est qu'un intermédiaire pour les calculs.

Dans les directions ( $d = \text{NOo}/\text{SEe}$ ,  $d = \text{NO n}/\text{SEs}$ ,  $d = \text{NE n}/\text{SOs}$ ,  $d = \text{NEe}/\text{SOo}$ ) la variance conditionnelle normalisée (i.e. ramenée au réseau de pas 1) est :

$$\begin{aligned}\sigma_{d_{norm}}^2 &= g(\lambda) * \sigma_d^2 \quad \text{où} \\ g(\lambda) &= \frac{\frac{P_{11}(\lambda)}{(2 + \lambda_{11}(d))}}{\frac{P_{27}(\lambda)}{(2 + \lambda_{27}(d))}}\end{aligned}$$

Dans les directions ( $d = \text{NE/SO}$ ,  $d = \text{NO/SE}$ ) la variance conditionnelle normalisée (i.e. ramenée au réseau de pas 1) est :

$$\sigma_{d_{norm}}^2 = h(\lambda) * \sigma_d^2 \quad \text{où}$$

$$h(\lambda) = \frac{\frac{P_{11}(\lambda)}{(2+\lambda_{11}(d))}}{\frac{P_{16}(\lambda)}{(2+\lambda_{16}(d))}}$$

Les 2 fonctions  $h(\lambda)$  et  $g(\lambda)$  sont tracées sur la figure 2.8.

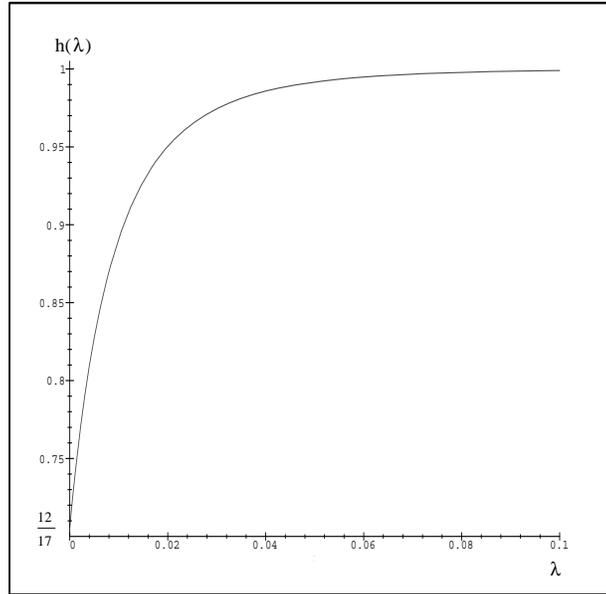
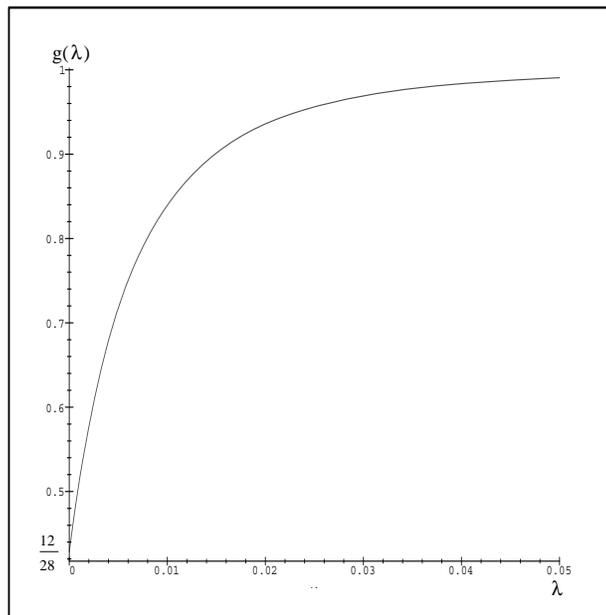
Suivant la direction considérée, nous estimons soit  $\lambda_{16}$ , soit  $\lambda_{27}$ . Puis, nous estimons  $\lambda$  à partir des équations (2.13), ou (2.14). Finalement, nous normalisons les variances conditionnelles locales par  $h(\lambda)$  ou  $g(\lambda)$  suivant la direction considérée.

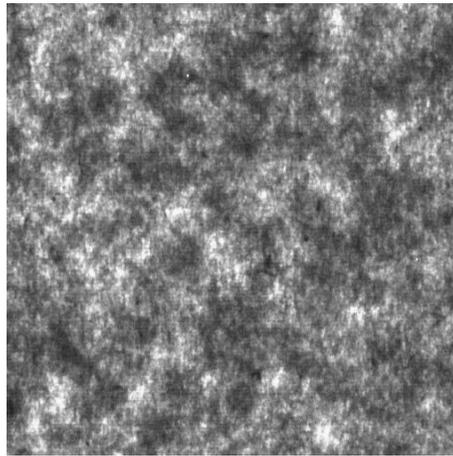
L'algorithme est le suivant :

#### Algorithme 2.4.1 (Estimation des paramètres de texture)

- ① *Nous estimons les variances conditionnelles locales dans toutes les directions à partir des données.*
- ② *Nous estimons  $\lambda_{16}$  ou  $\lambda_{27}$  suivant la direction considérée.*
- ③ *Nous estimons  $\lambda$  à partir des équations (2.13) et (2.14).*
- ④ *Nous normalisons les variances conditionnelles par  $h(\lambda)$  dans les directions NE/SO et NO/SE et par  $g(\lambda)$  dans les directions NOo/SEe, NO<sub>n</sub>/SEs, NE<sub>n</sub>/SOs et NEe/SOo.*

Afin de voir l'apport d'une telle correction, nous avons analysé la distribution des variances conditionnelles locales estimées sur une image de texture isotrope (Figure 2.9). Les courbes représentant les fonctions de répartition de la variance conditionnelle  $\sigma_d^2$  après normalisation des paramètres estimés (Figure 2.10 (b)) sont plus proches les unes des autres qu'avant (Figure 2.10 (a)). En d'autres termes, les distances (au sens de la distance de Kolmogorov-Smirnov) entre les distributions sont plus petites après avoir corrigé les paramètres. Ces observations montrent la nécessité d'une telle correction. Soulignons que dans l'étude précédente nous avons associé les pixels à des mesures ponctuelles.

(a): fonction  $h(\lambda)$ (b): fonction  $g(\lambda)$ **Figure 2.8:** Fonctions de correction  $h(\lambda)$  et  $g(\lambda)$ .



**Figure 2.9:** *Texture isotrope (512×512) : surface d'une feuille de papier.*

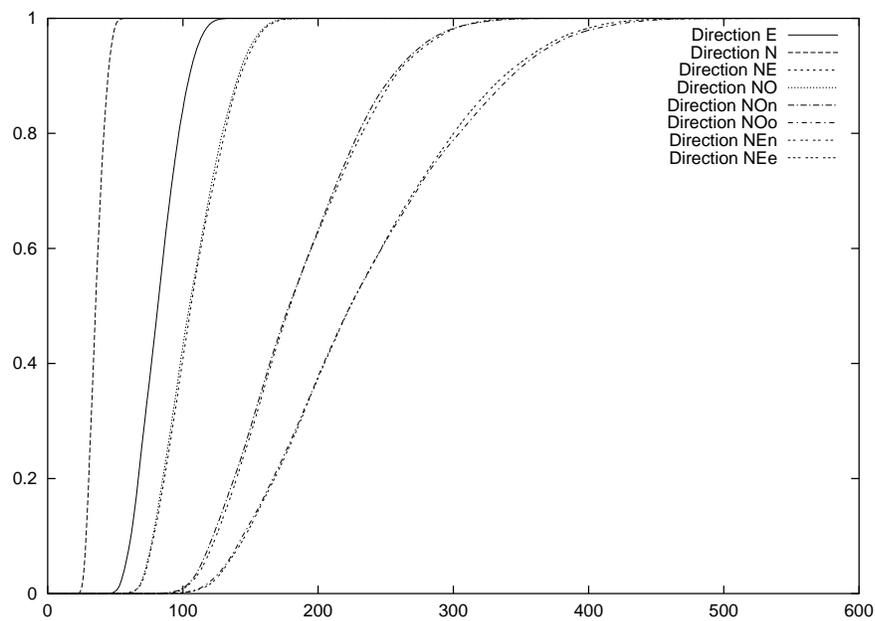
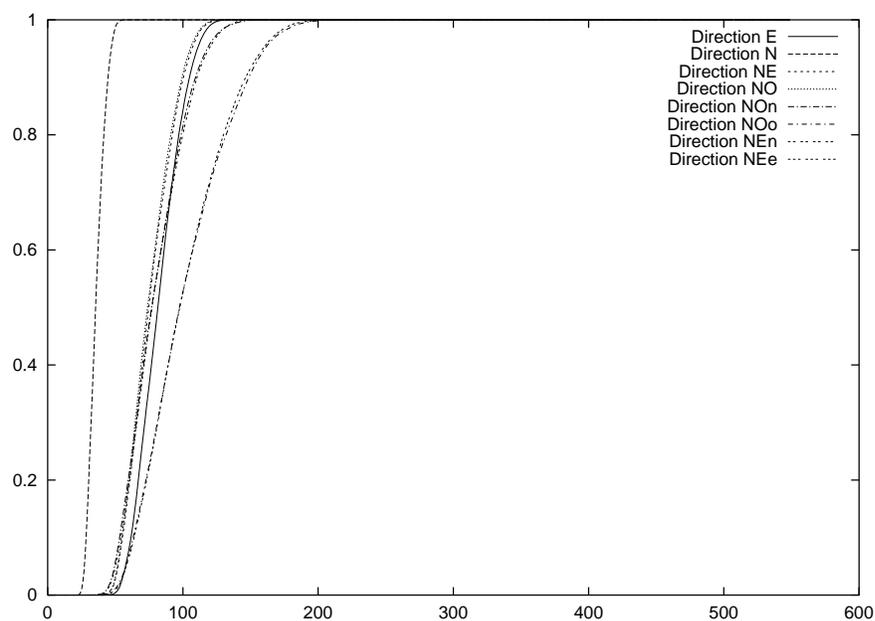
#### 2.4.4 Résultats

Dans le paragraphe 2.3.3, nous avons expliqué pourquoi nous avons besoin de tenir compte des différentes directions. Dans ce paragraphe, nous présentons des résultats complémentaires. Nous avons donc estimé sur l'image 2.5 les variances conditionnelles locales pour chacun des huit modèles.

Les résultats sont présentés sur les figures 2.11 (a)(b)(c)(d)(e)(f) et 2.12 (g)(h). Nous remarquons que la serre n'apparaît pas sur l'image 2.12 (h). Cette image représente l'estimée du paramètre de texture du modèle défini dans la direction NOo, direction dans laquelle est orientée la serre. D'autre part, elle répond comme la zone urbaine, très nettement sur l'image 2.11 (e) et nettement sur l'image 2.12 (g). Ces images représentent l'estimée du paramètre de texture des modèles définis dans les directions N et NO, directions dans lesquelles l'objet présente une variation plus forte des niveaux de gris. L'approche qui consiste à estimer huit paramètres à partir de huit modèles différents, permet donc de distinguer les objets qui répondent fortement quelque soit la direction (e.g. les ville), les objets qui répondent faiblement quelque soit la direction (e.g. les champs) et les objets qui répondent faiblement ou fortement suivant le modèle considéré (e.g. les serres). Soulignons, en outre, que la serre apparaît également nettement sur l'image 2.12 (j). Cette image représente l'estimée du paramètre de texture du modèle isotrope 2D décrit dans le paragraphe 2.3. Ce paramètre, s'il permet de distinguer la zone urbaine et champs, ne permet pas de distinguer la zone urbaine et la serre.

Ceci confirme l'intérêt d'une analyse directionnelle.

Cependant, notre objectif n'est pas de savoir dans quelle direction est orienté

(a): Avant la normalisation (CPU=341.32s, fenêtre  $15 \times 15$ )(b): Après la renormalisation par décimation (CPU=469.18s, fenêtre  $15 \times 15$ )**Figure 2.10:** Fonctions de répartition du paramètre de texture estimé dans chacune des 8 directions.

tel ou tel objet. C'est pourquoi nous combinons les paramètres estimés dans les huit directions afin d'obtenir un paramètre discriminant, robuste par rapport à la grande variété des images à traiter (type de capteur différent, résolution différente, et structure même de la zone urbaine différente). Pour cela, nous avons réalisé des tests sur des zones diverses (ville, champs, objets orientés) extraites d'images SPOT3, de simulations SPOT5, d'images ERS1.

En chaque pixel, nous estimons la variance conditionnelle à partir de la méthode des queues de comète et ceci dans huit directions. Les figures 2.13, 2.14, 2.15, 2.16 représentent pour une zone extraite d'une image SPOT3 ou d'une simulation SPOT5, la rose des directions des paramètres estimés. De même la figure 2.17 représente pour une zone extraite d'une image ERS1, la rose des directions des paramètres estimés. Pour chacune des huit directions le rayon représente la valeur du paramètre estimé. Cette rose est invariante par rotation d'un angle  $\pi$ . La représentation des résultats sous cette forme nous permet de faire les remarques qui suivent.

La plupart des milieux urbains sont caractérisés par une forte variance dans la plupart des directions (Figures 2.15, 2.16, 2.17 bas). Par ailleurs les zones de champs (Figure 2.13, 2.17 bas), de forêt ou encore de mer présentent une faible variance dans toutes les directions comparativement à la zone de ville sur le même type d'images (optique ou radar). Enfin des constructions réalisées par l'homme telles que les serres, les hangars ou encore les routes, de même que les petits parcellaires ont une orientation privilégiée (Figure 2.14) dans laquelle la variance conditionnelle estimée est plus faible.

Nous avons envisagé dans un premier temps de ne conserver pour chaque pixel que la réponse la plus faible sur les huit calculées. Cependant, certains des pixels situés à la lisière des villes étaient mal classifiés. En effet, la fenêtre centrée sur un de ces pixels et utilisée pour estimer les variances conditionnelles contient un mélange de deux textures (urbain et non urbain). L'heuristique (méthode des queues de comètes) utilisée pour estimer les paramètres dans ce cas de non stationnarité (moyenne locale constante par morceaux) suppose que les deux parties de la matrice des probabilités conditionnelles puissent être différenciées. Ceci n'est pas toujours le cas. Toutefois, nous conservons quand même la valeur minimale qui nous servira par la suite à créer une image de marqueurs afin d'éliminer les fausses alarmes (cf. paragraphe 3.2.3).

Une grande partie de cette ambiguïté a été levée en ne considérant pas les valeurs extrêmes. Pour chaque pixel, nous classons les huit réponses dans l'ordre

croissant. Les valeurs extrêmes, peu caractéristiques, sont éliminées. Finalement, le paramètre de texture est une combinaison linéaire des deux valeurs centrales obtenue par ACP (Analyse en Composantes Principales). Nous avons également fait des tests en conservant les 4 ou 6 valeurs centrales et en les combinant par une ACP, ce qui a pour conséquence une dégradation du résultat.

Pour la plupart des points de la ville répondant fortement dans toutes les directions, la valeur du paramètre proposé reste élevée. Pour les champs, répondant faiblement dans toutes les directions, cette valeur reste faible. Quant aux points ambigus, la plupart se retrouvent bien classés. En effet, les pixels "ville" qui répondent très faiblement dans une, deux ou même trois directions mais fortement dans les autres seront alors bien classés, i.e. dans la classe zone urbaine. La valeur du paramètre associée à ces pixels est forte comme en plein centre ville.

D'autre part, nous avons également remarqué que ce paramètre donnait des régions beaucoup plus homogènes que le paramètre de température isotrope (Figures 2.12 (j), 2.18 (b)). Sur l'image du paramètre proposé (Figures 2.12 (i), 2.18(c)) la zone urbaine est plus lisse que sur l'image du paramètre isotrope qui présente des "trous". Nous l'avons également vérifié sur d'autres images.

A partir d'une image en niveaux de gris, nous estimons donc une image de paramètre (Figures 2.12 (i), 2.18 (c)) sur laquelle la ville est caractérisée par de fortes valeurs<sup>6</sup>.

## 2.4.5 Comparaison avec deux méthodes classiques

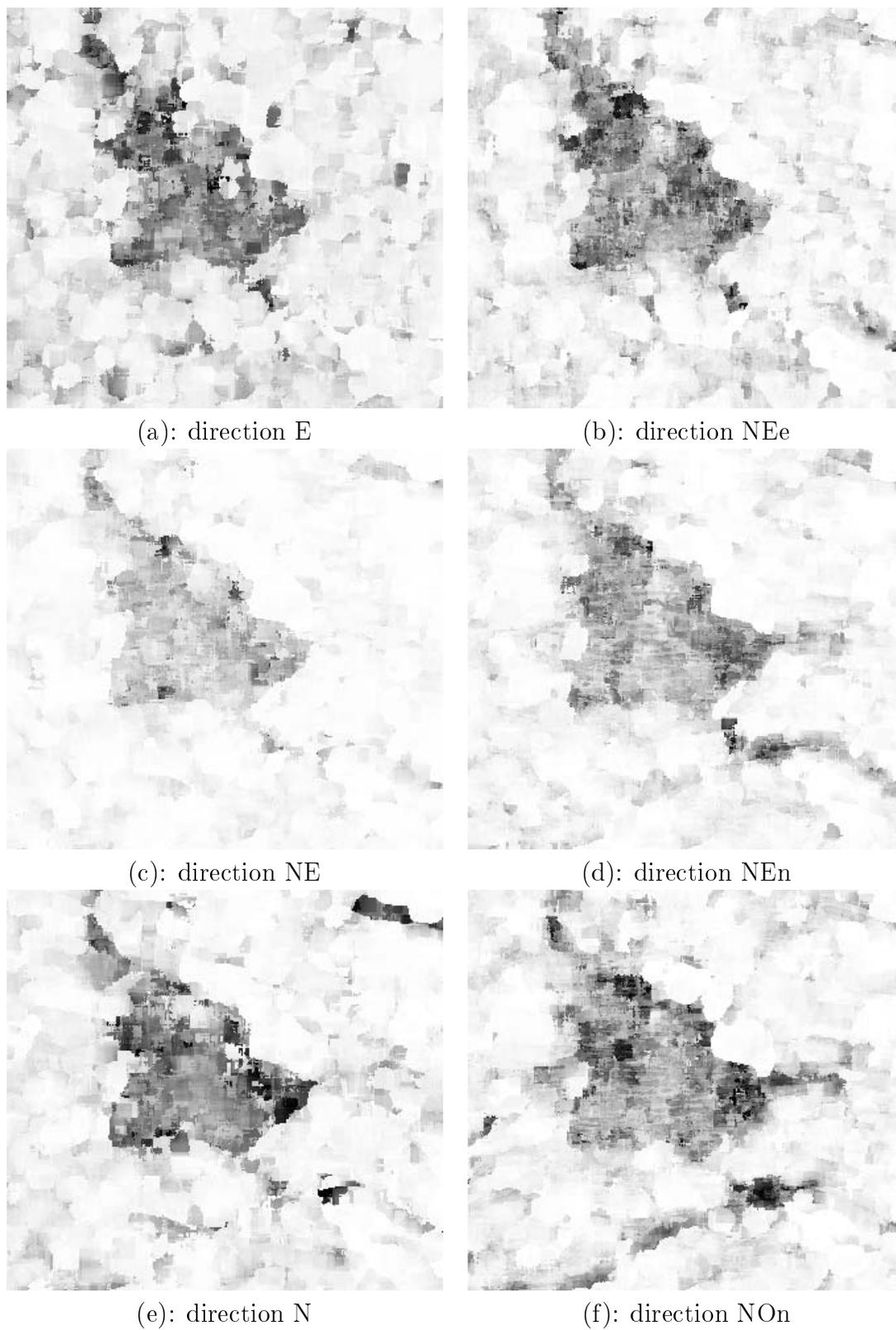
Dans le chapitre 1.2.3 nous avons décrit de nombreuses méthodes d'analyse de la texture proposées dans la littérature. Nous comparons notre méthode avec deux méthodes anisotropes classiques : l'analyse par banc de filtres de Gabor [63] et l'analyse à partir de matrices de cooccurrence [22].

### 2.4.5.1 Filtres de Gabor

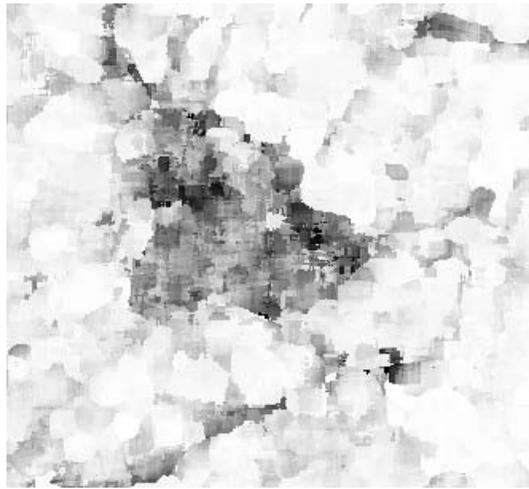
L'analyse proposée fait partie des approches par filtrage multicanal. Dans le cas présent les différents canaux sont représentés par un ensemble de filtres de Gabor. Dans [36], les auteurs utilisent des filtres en quadrature. Nous utilisons des filtres réels, pairs et symétriques comme dans [63]. Nous avons vérifié que les

---

<sup>6</sup>Les images présentée sont en vidéo inversée, i.e. les pixels sombres ont une valeur de paramètre élevée.



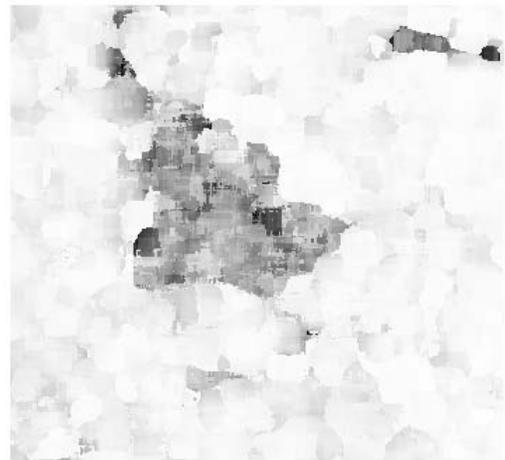
**Figure 2.11:** Images des estimées (fenêtre de taille  $17 \times 17$ ) des paramètres de texture (vidéo inversée) sur l'image de Toreilles.



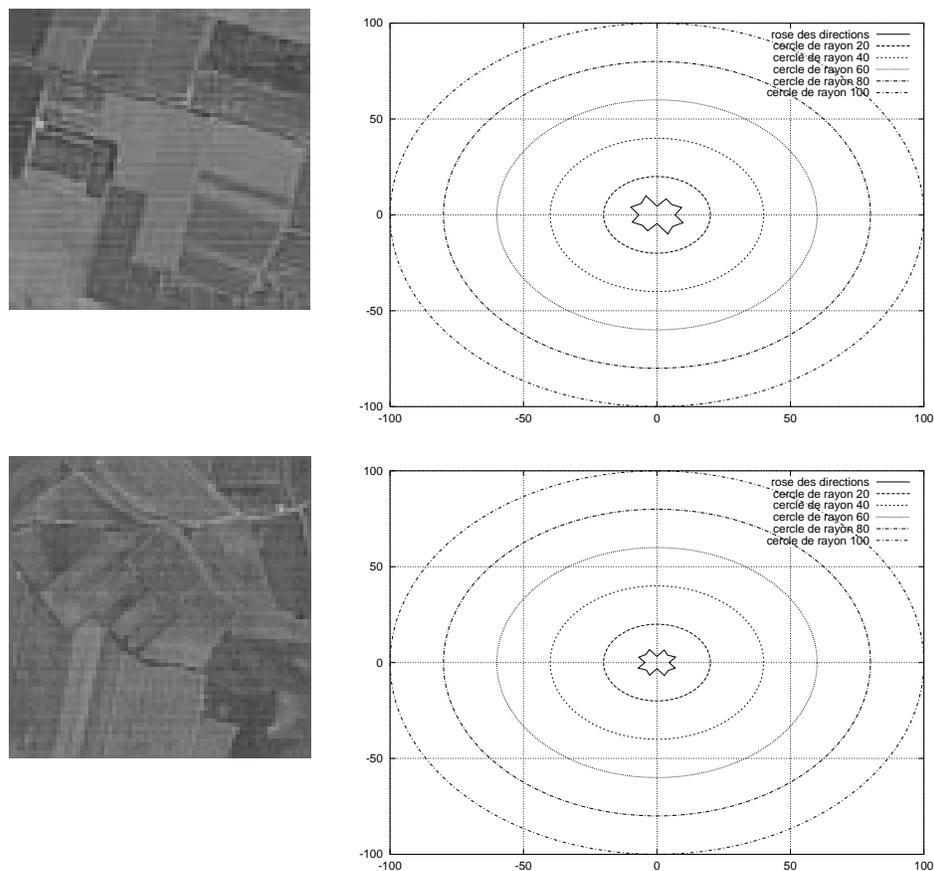
(g): direction NO



(h): direction NOo

(i): Paramètre global  
(combinaison linéaire des valeurs centrales)(j): Paramètre du modèle  
isotrope 2D

**Figure 2.12:** Images des estimées (fenêtre de taille  $17 \times 17$ ) des paramètres de texture (vidéo inversée) sur l'image de Toreilles.



**Figure 2.13:** *Rose des directions sur des images de champs extraites d'une simulation SPOT5 (en haut) et d'une image SPOT3 (en bas).*

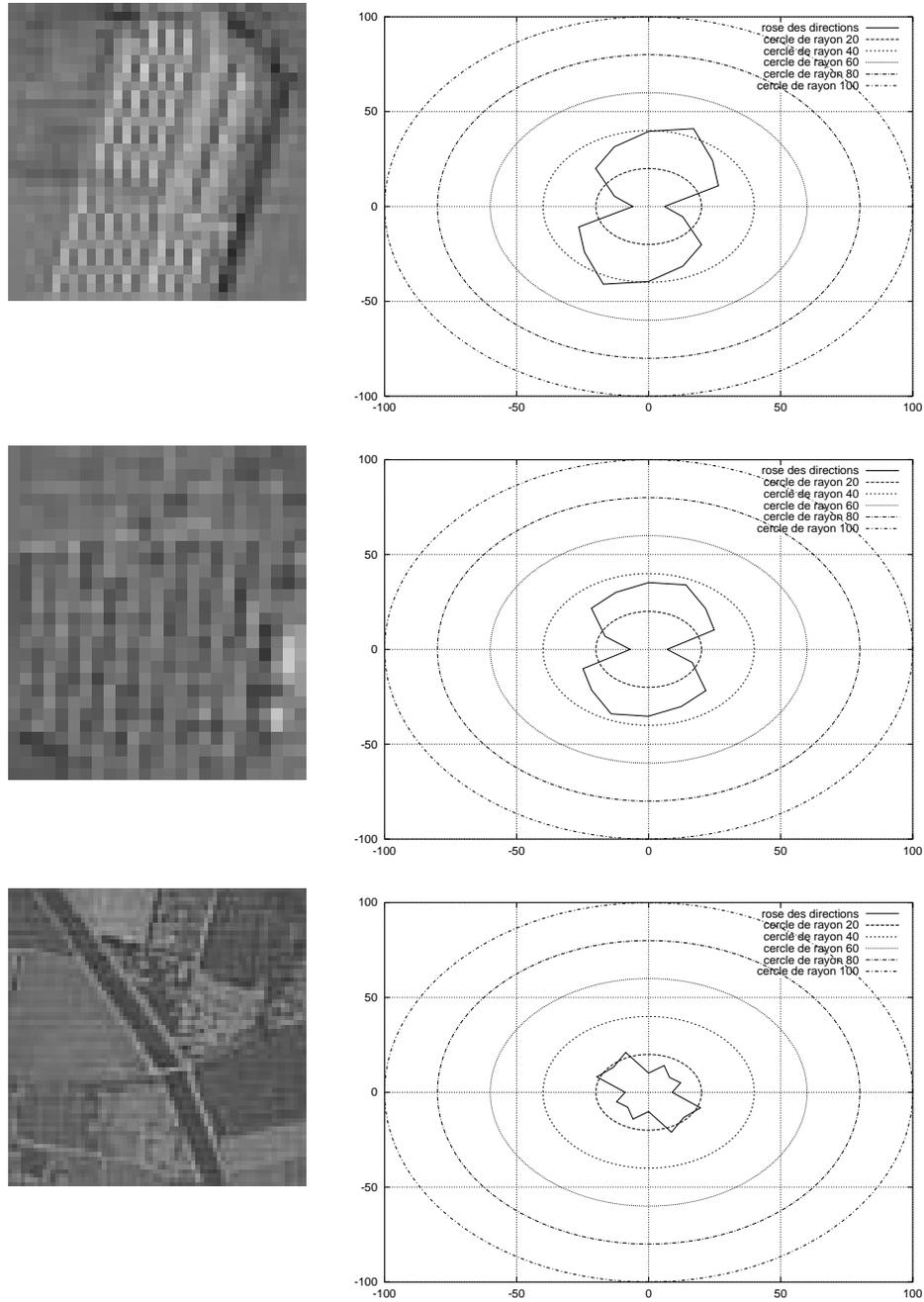
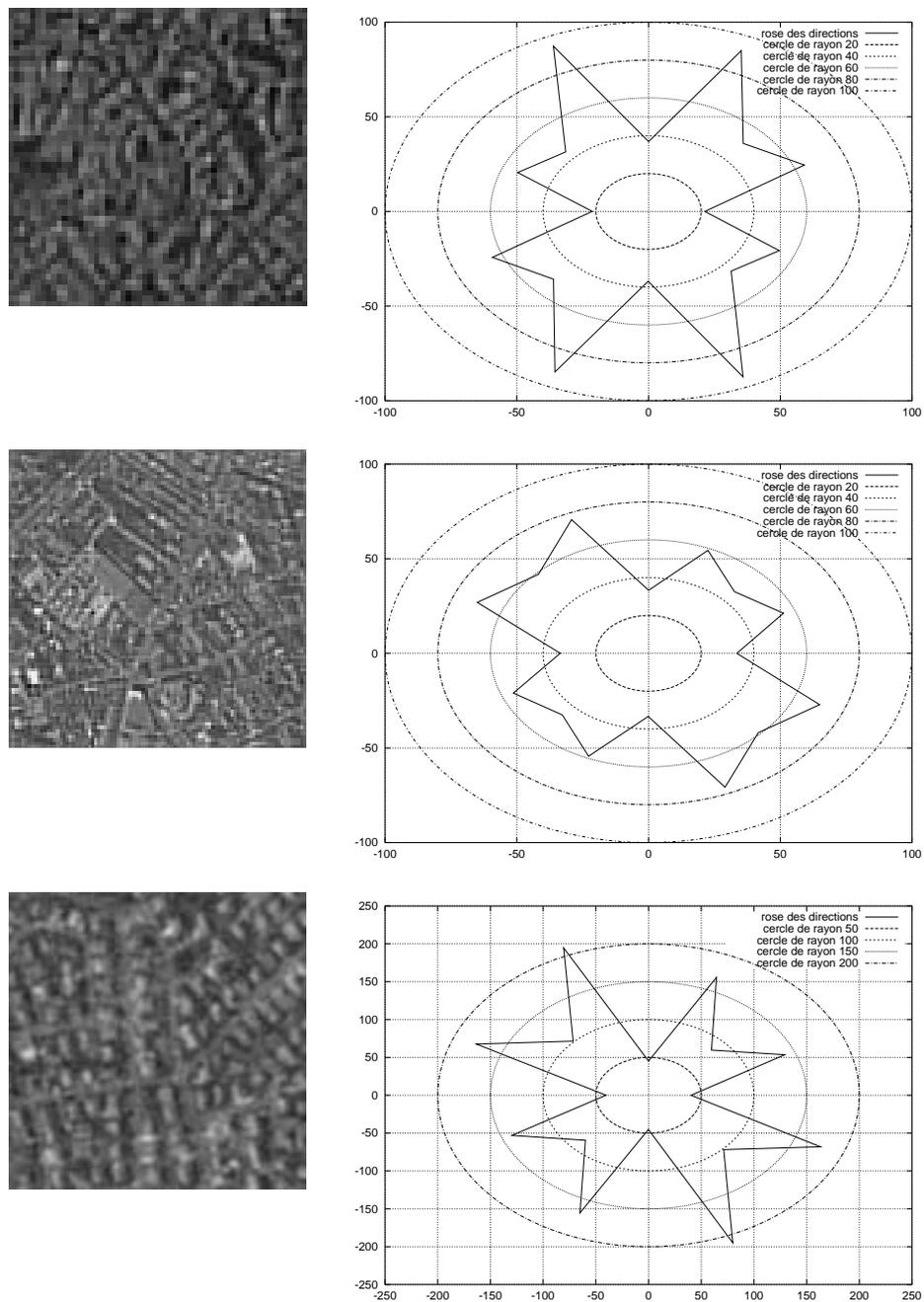
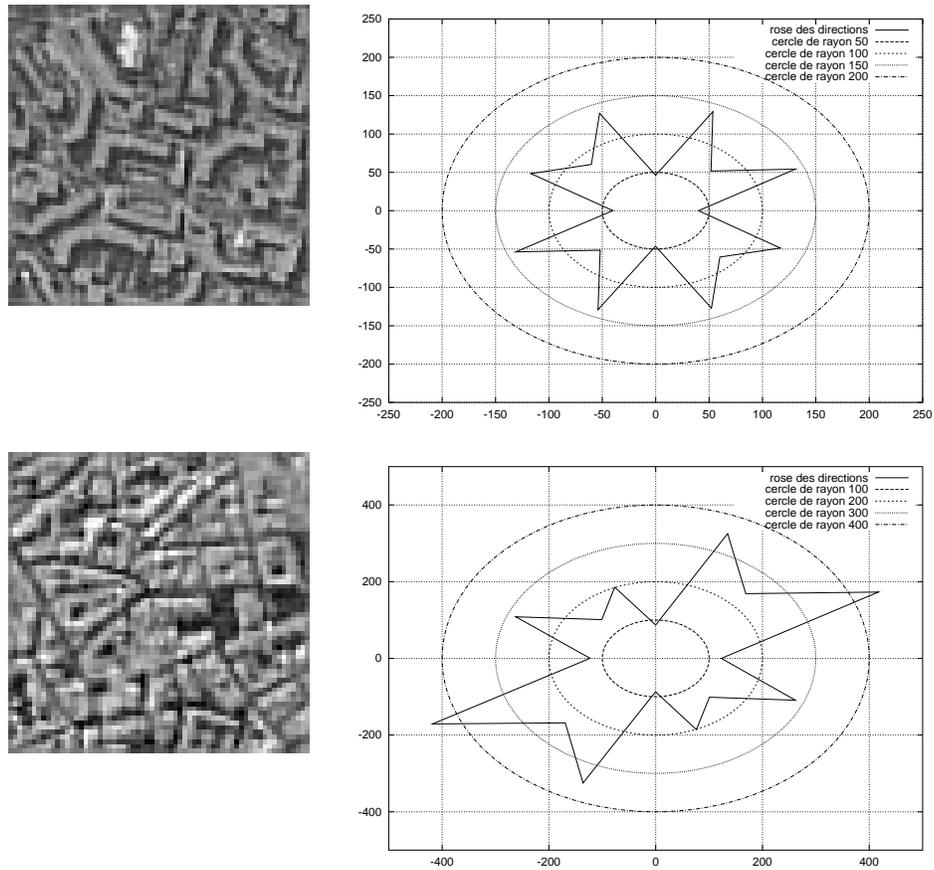


Figure 2.14: *Rose des directions sur des images ayant une orientation privilégiée extraites d'une simulation SPOT5.*



**Figure 2.15:** *Rose des directions sur des images de zones urbaines extraites d'une images SPOT3 (image du haut) et de simulations SPOT5 (les deux images du bas).*



**Figure 2.16:** *Rose des directions sur des images de zones urbaines extraites de simulations SPOT5.*

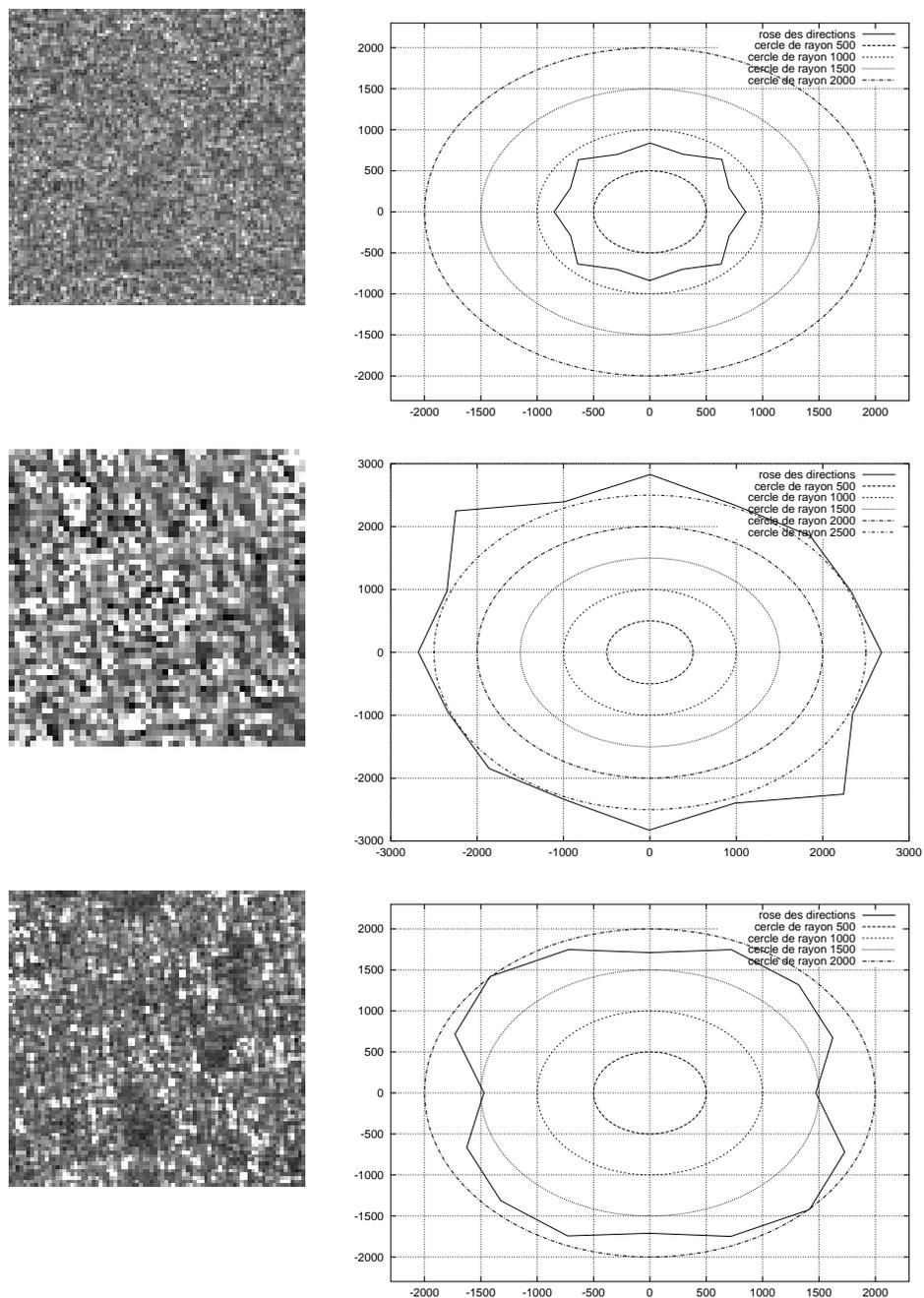
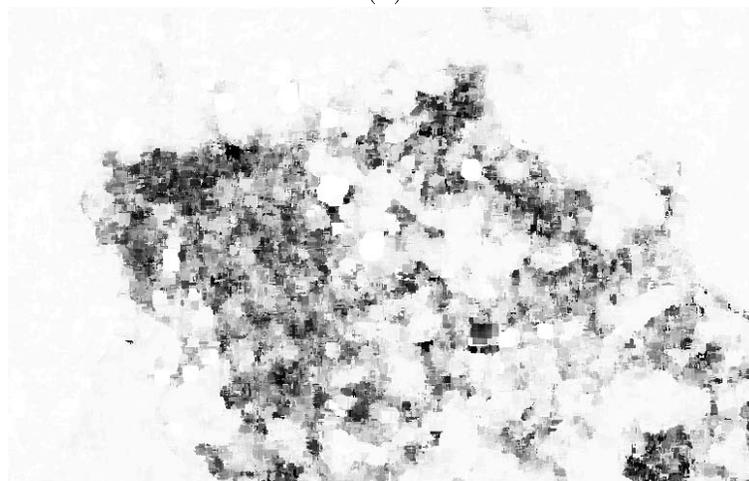


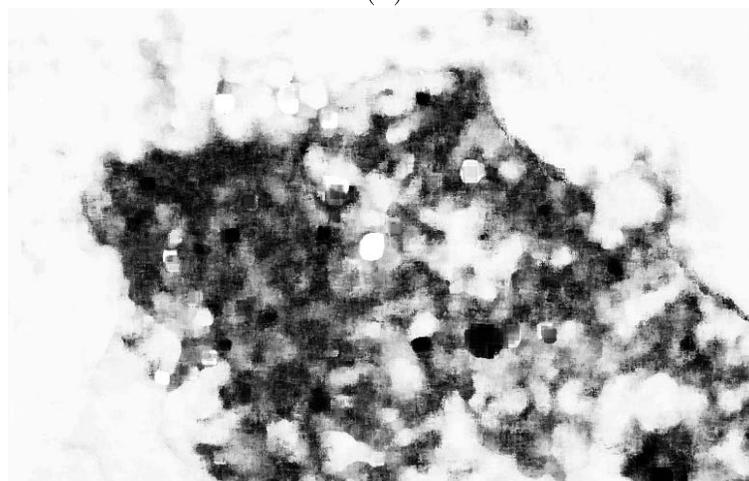
Figure 2.17: Rose des directions sur des images extraites d'images ERS1 (haut : champs, les 2 images du bas : ville).



(a)



(b)



(c)

**Figure 2.18:** (a): Image SPOT3 de la ville de Cayenne, (b): Estimée de la variance conditionnelle locale du modèle isotrope 2D, (c): Estimée du paramètre de texture proposé.

résultats obtenus sont de meilleure qualité (en l'occurrence plus discriminants) si nous n'utilisons que les filtres pairs. Dans [85], les auteurs justifient ce choix à partir de considérations psychophysiques.

La méthode consiste donc à convoluer l'image en niveaux de gris avec une série de filtres de Gabor 2D et à calculer l'énergie locale pour chaque pixel (Figure 2.19). Chaque filtre de Gabor  $G$  est une onde plane sinusoïdale qui est modulée par une enveloppe gaussienne. Le filtre de Gabor pair, symétrique est défini dans la direction  $\theta = 0^\circ$  de la manière suivante :

$$G(s_x, s_y/\theta = 0^\circ, \lambda, \phi = 0^\circ) = \exp \frac{-(s_x^2 + s_y^2)}{2\sigma^2} \cos\left(\frac{2\pi}{\lambda}s_x^2 + \phi\right) \quad (2.18)$$

$(s_x, s_y)$  sont les coordonnées du pixel  $s$  dans le système de coordonnées que nous notons  $(X, Y)$ .

Le filtre défini dans une direction  $\theta$  (angle avec l'axe X) quelconque est obtenu par rotation du système de coordonnées  $(X, Y)$  d'un angle  $\theta$ . Notons que le filtre impair est obtenu en posant  $\phi = \frac{\pi}{2}$ . Ainsi de nombreux filtres peuvent être générés en faisant varier la longueur d'onde  $\lambda$ , l'orientation  $\theta$ , et l'écart type de la fenêtre gaussienne  $\sigma$ .

Après avoir convolué l'image initiale par un banc de filtres, nous calculons pour chaque pixel  $t(t_x, t_y)$  sur une fenêtre  $W_{t_x, t_y}$  centrée en  $t$  de taille  $M$  une énergie locale définie comme suit :

$$\mathcal{E}_G(t_x, t_y/\theta, \lambda) = \frac{1}{M^2} \sum_{(s_x, s_y) \in W_{t_x, t_y}} |G(s_x, s_y/\theta, \lambda, \phi)|^2 \quad (2.19)$$

Afin de comparer cette méthode avec la nôtre, nous avons choisi les mêmes 8 orientations pour  $\theta$  (Figure 2.6) que pour nos huit modèles, i.e.  $\theta \in \{0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}\}$ . Soulignons que ces filtres sont sélectifs du point de vue de l'orientation et de la fréquence.

Dans les tests qui suivent  $\sigma = 12$  pour toute les directions sauf les directions  $E$  et  $N$  pour lesquelles  $\sigma = 9$ , la taille du filtre est  $25 \times 25$ , et  $M = 11$ . Ces paramètres ont été déterminés empiriquement, sachant que  $\sigma$  dépend de l'échelle de l'image d'entrée. La valeur de  $\lambda$  peut varier suivant l'orientation, c'est pourquoi nous la précisons dans le titre des images 2.21 et 2.22. Les huit filtres utilisés sont présentés figure 2.20.

Les bords des zones sont très flous (Figures 2.21 et 2.22) comparés aux images présentées figures 2.11 et 2.12. En effet, associés à nos modèles nous avons une

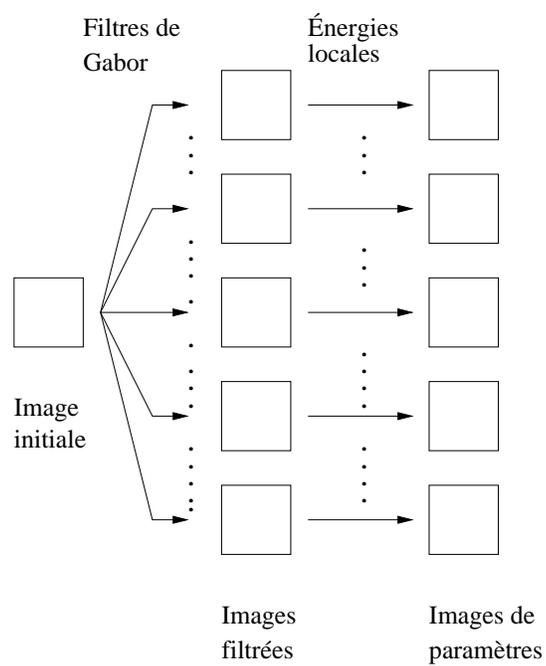


Figure 2.19: Estimation des paramètres de texture

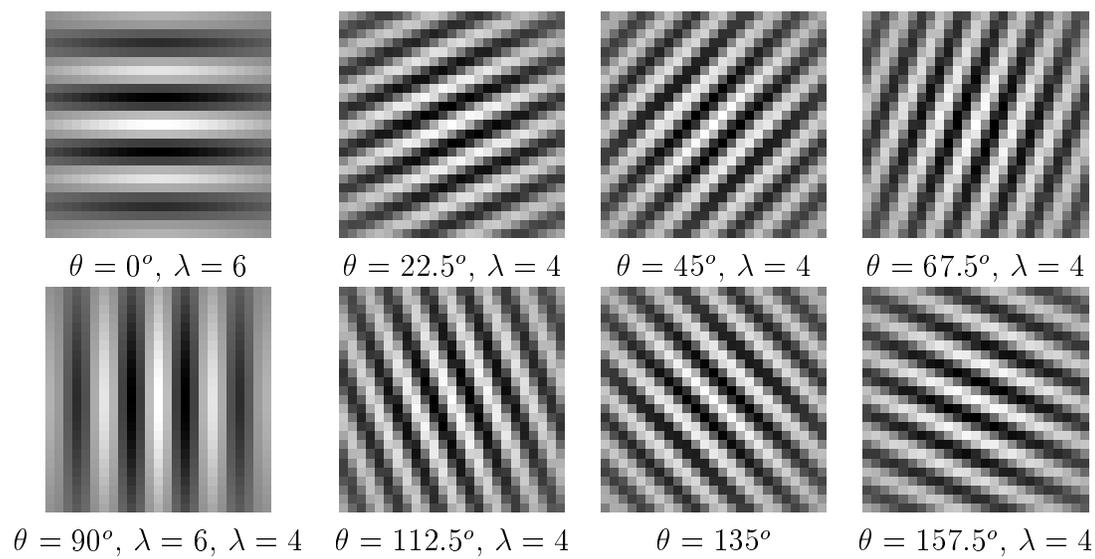


Figure 2.20: Filtres de Gabor

méthode d'estimation et une heuristique qui nous permettent d'avoir des bords francs entre les zones.

Pour les filtres de Gabor, hormis la taille de la fenêtre et les directions, on doit également fixer la longueur d'onde  $\lambda$  et la variance de l'enveloppe gaussienne  $\sigma$ . Il est avantageux lorsque l'on cherche à automatiser un processus d'avoir un faible nombre de paramètres à estimer ou à fixer empiriquement, ce qui est le cas pour l'approche que nous avons proposée.

L'avantage de l'approche par filtres de Gabor est un temps de calcul (cf. Tableau 2.1) plus faible qu'avec notre approche.

### 2.4.5.2 Matrices de cooccurrence

La matrice de cooccurrence  $M_{\mathbf{d}} = [m_{gg'}]$  est construite en estimant les probabilités  $m_{gg'}$  de passer d'un niveau de gris  $g$  à un niveau de gris  $g'$  en se déplaçant d'un vecteur  $\mathbf{d}(\mathbf{d}_x, \mathbf{d}_y)$  sur le réseau des sites :

$$m_{gg'} = P[X_s = g, X_t = g' | t_x = s_x + \mathbf{d}_x, t_y = s_y + \mathbf{d}_y]$$

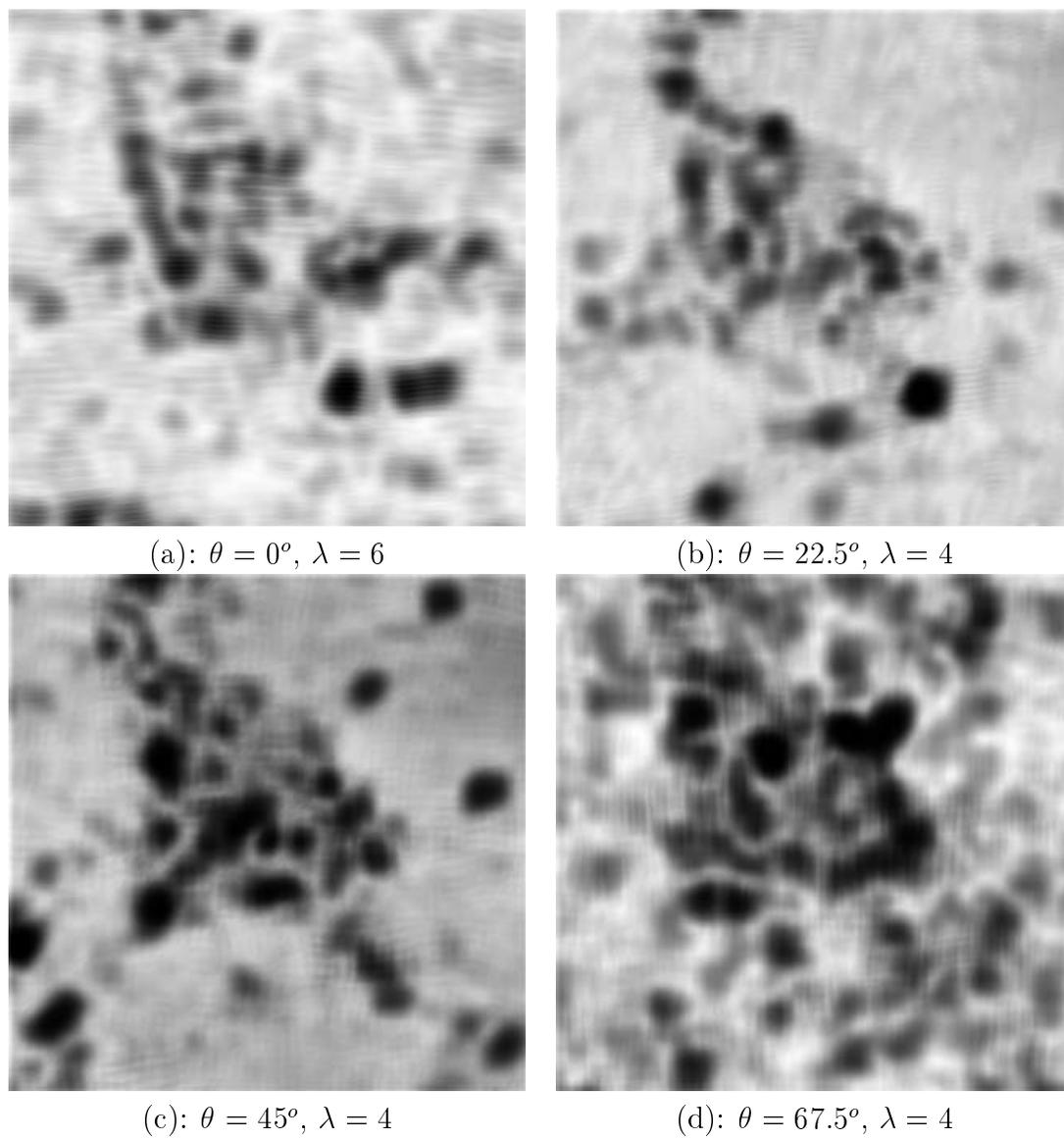
$X_s$  est le niveau de gris du pixel  $s$  de coordonnées  $(s_x, s_y)$ .

$X_t$  est le niveau de gris du pixel  $t$  de coordonnées  $(t_x, t_y)$ .

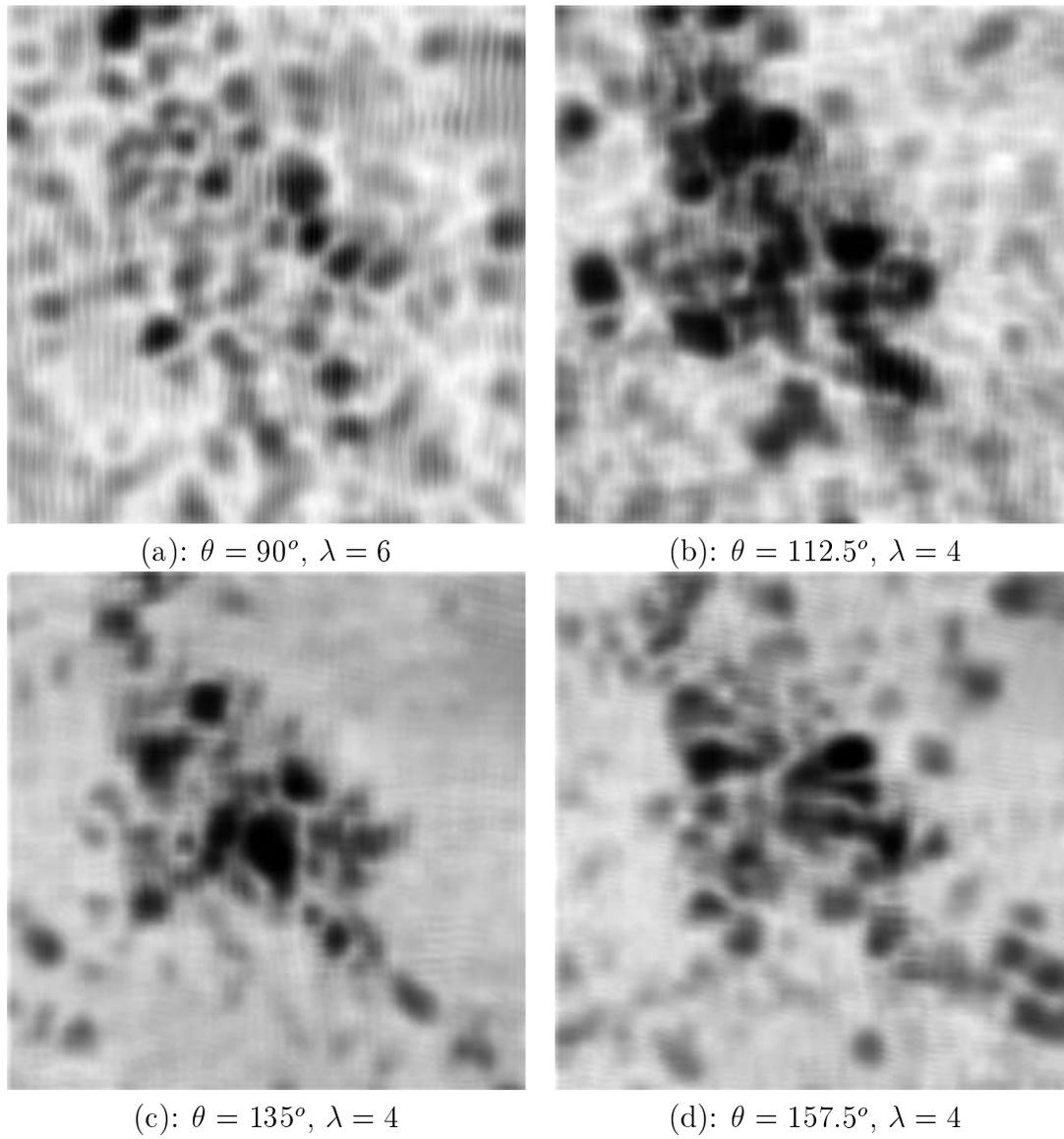
Ainsi à chaque choix de vecteur de déplacement  $\mathbf{d}$  correspond une matrice de cooccurrence. Une matrice est estimée sur une fenêtre glissante de taille  $9 \times 9$  centrée sur le pixel auquel elle est associée.

Une fois la matrice construite, nous extrayons des paramètres. Les paramètres que nous avons utilisés sont ceux proposés par Connors et al. dans [22]. De même que pour les matrices de miscibilité, nous ramenons la dynamique de l'image à 32 niveaux de gris. Nous choisissons 8 vecteurs de déplacement  $\mathbf{d}$  afin de prendre en compte les mêmes orientations que dans notre modèle. Nous ajoutons à la matrice  $M_{\mathbf{d}}$  sa transposée et nous divisons chaque élément  $m_{gg'}$  par 2 afin de rendre  $M_{\mathbf{d}}$  symétrique, i.e. invariante par rotation de  $\mathbf{d}$  d'un angle  $\pi$ . En effet, nous avons précisé (Figure 2.6) que les directions NOo et SEe étaient les mêmes.

Les 5 paramètres utilisés sont les suivants :



**Figure 2.21:** *Énergies de Gabor calculées à la sortie de filtres orientés (vidéo inversée).*



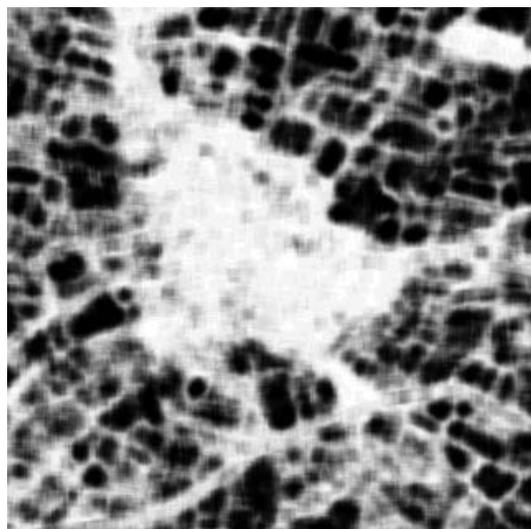
**Figure 2.22:** *Énergies de Gabor calculées à la sortie de filtres orientés (vidéo inversée).*

$$\begin{aligned}
1. \text{ Énergie} &= \sum_{gg'} m_{gg'}^2 \\
2. \text{ Inertie} &= \sum_{gg'} (g - g')^2 m_{gg'} \\
3. \text{ Cluster shade} &= \sum_{gg'} ((g - \mu_g) + (g' - \mu_{g'}))^3 m_{gg'} \\
4. \text{ Cluster prominence} &= \sum_{gg'} ((g - \mu_g) + (g' - \mu_{g'}))^4 m_{gg'} \\
5. \text{ Inverse different moments} &= \frac{1}{1 + (g - g')^2} m_{gg'} \\
\text{où } \mu_g &= \sum_g g \sum_{g'} m_{gg'} \\
\text{où } \mu_{g'} &= \sum_{g'} g' \sum_g m_{gg'}
\end{aligned}$$

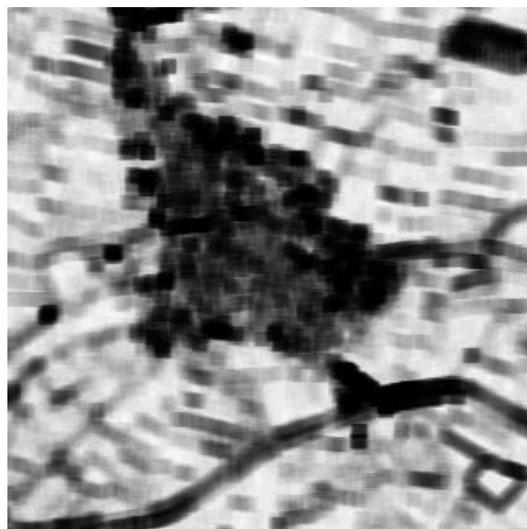
Cette approche est moins souple que les précédentes car le choix des paramètres à extraire de  $M_d$  n'est pas évident. D'autre part, nous devons réduire la dynamique des niveaux de gris pour stocker des matrices de taille plus faible. Sur les images de paramètre, le fenêtrage est très visible. Les bords sont imprécis. Enfin, pour une direction donnée nous estimons plusieurs paramètres, un seul n'étant pas assez caractéristique. Dans les résultats présentés, nous donnons pour une direction donnée (Figure 2.23) les 5 paramètres qui sont les plus discriminants. Par rapport aux méthodes précédentes celle-ci est plus "lourde" car le nombre d'images de paramètres à traiter est de 40 au lieu de 8 dans les deux cas précédents. D'autre part, elle est plus coûteuse en temps de calcul (Tableau 2.1).

	Var. Cond.	Gabor	GLCM
CPU(s.)	102	13	192/direction

**Tableau 2.1:** Temps de calcul pour chacune des 3 méthodes sur station SUN ultra1 (124MB, 167 MHz).



(a): Énergie



(b): Inertie

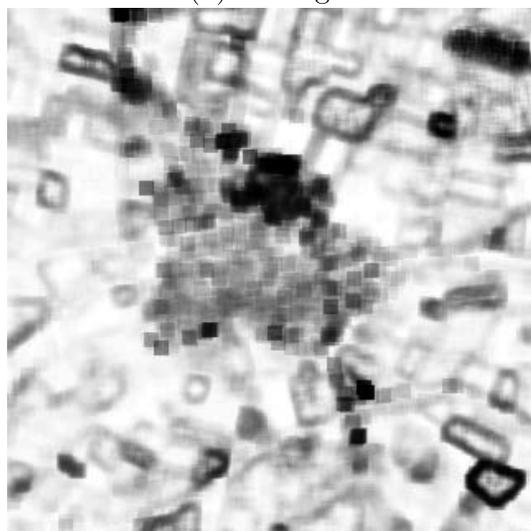
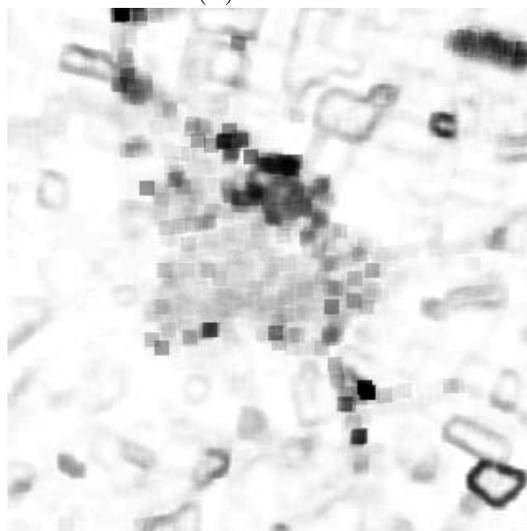
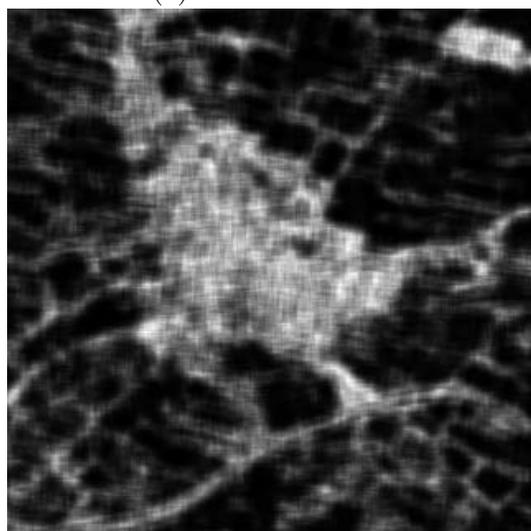
(c): *cluster shade*(d): *cluster prominence*(e): *Inverse different moments*

Figure 2.23: Paramètres estimés dans la direction  $N$ , i.e.  $\mathbf{d} = (0, -1)$  (vidéo inversée).

### 2.4.6 Conclusion

Les analyses faites ont montré l'utilité de l'analyse anisotrope de la texture pour caractériser l'urbain. L'approche que nous avons proposée présente plusieurs avantages. Les modèles définis sont simples. Il est facile d'introduire de nouvelles directions en conservant la méthode d'estimation. La difficulté peut venir alors de la phase de normalisation des paramètres.

Par rapport aux autres méthodes classiques décrites ci-dessus (filtres de Gabor, matrices de cooccurrence), la nôtre est plus souple. En effet, les seuls paramètres à fixer sont la taille de la fenêtre et les directions à prendre en compte.

D'autre part comme notre analyse de texture est fondée sur des modèles, nous avons des méthodes pour estimer ces paramètres et également des heuristiques pour contourner certains problèmes de mélanges de textures. Ceci nous permet d'obtenir des images de paramètre avec des bords nets entre les zones. Enfin, la variance conditionnelle locale est un paramètre qui est indépendant de la moyenne locale.

En conclusion, l'approche proposée nous permet d'obtenir un paramètre caractéristique de la zone urbaine. Pour cela nous utilisons 8 modèles markoviens basés sur des chaînes. Le paramètre, i.e. la variance conditionnelle locale, est estimé de manière rapide par la méthode des "queues de comètes". Par ailleurs, comme nous le verrons par la suite, ce paramètre est robuste par rapport à la résolution des images et au type de capteur.

## DANS CE CHAPITRE:

3.1	Classification floue . . . . .	72
3.1.1	Ensembles flous . . . . .	72
3.1.2	Critère classique . . . . .	72
3.1.3	Critère entropique . . . . .	74
3.1.4	Résultats expérimentaux . . . . .	79
3.2	Régularisation par modélisation markovienne . . . . .	81
3.2.1	Modèle de segmentation . . . . .	87
3.2.1.1	Critère du MAP . . . . .	87
3.2.1.2	Vraisemblance . . . . .	87
3.2.1.3	A priori . . . . .	89
3.2.2	Algorithme ICM ( <i>Iterated conditional modes</i> ) . . . . .	89
3.2.3	Image de marqueurs . . . . .	90
3.3	Résultats . . . . .	90
3.4	Conclusion . . . . .	94

# 3.

## Masque Urbain: Segmentation

Dans ce chapitre, nous proposons un nouvel algorithme de partition non supervisé prenant en compte un critère entropique. Cet algorithme nous permet de classifier l'image du paramètre de texture défini dans le chapitre précédent. Par ailleurs, il nous permet d'estimer les paramètres du modèle markovien utilisé pour la segmentation finale.

## 3.1 Classification floue

Dans l'étape précédente, nous avons transformé l'information de niveaux de gris en une information caractéristique de la texture urbaine. Dans cette seconde étape, nous utilisons cette information pour extraire le thème urbain.

Nous proposons un algorithme totalement non supervisé, i.e. le nombre de nuages de points (ou *clusters* en anglais) n'est pas fixé a priori. Pour cela nous modifions le critère classique du *fuzzy Cmeans* [71] en ajoutant un terme d'entropie. Le nombre optimal de nuages de points est ainsi déterminé de manière automatique.

### 3.1.1 Ensembles flous

Nous travaillons dans le cadre de la partition floue, des ensembles flous tels qu'ils ont été définis par Zadeh [132].

**Définition 3.1.1** *Un ensemble flou  $A$  est caractérisé par une fonction  $f_A$  qui varie continuellement sur  $[0, 1]$  et qui associe à chaque objet  $x$  d'un ensemble  $X$  la valeur  $f_A(x) \in \mathbb{R}$  qui représente le degré d'appartenance de  $x$  à l'ensemble  $A$ . Plus généralement, le domaine de définition de  $A$  peut être réduit à un sous ensemble de  $X$ .*

Pour notre problème, les ensembles flous sont les nuages de points et les objets sont les pixels. Une partition de l'image est alors caractérisée par une matrice de partition  $\mathbf{U} = [u_{ij}]$ , où  $u_{ij}$  caractérise le degré d'appartenance du pixel  $j$  au nuage de points  $i$  [71].

Une méthode visant à trouver des groupements naturels dans l'espace d'état est alors mise en place afin de déterminer cette matrice de partition. Chaque pixel peut être caractérisé soit par son niveau de gris, soit par un paramètre de texture, soit encore par un vecteur de paramètres. L'avantage de cette méthode par rapport à la version "dure" est qu'elle nous permet de garder la majeure partie de l'information issue des données initiales pour la seconde partie de notre travail. En effet, à ce niveau de traitement, aucune décision définitive n'est prise quant à l'appartenance d'un pixel à tel ou tel nuage de points.

### 3.1.2 Critère classique

Dans l'algorithme classique du *fuzzy Cmeans*, décrit pour la première fois dans [15], le nombre de nuages de points est fixé de manière supervisée par l'opérateur.

L'étape de partition consiste alors à minimiser le critère suivant, qui est une généralisation au cas flou de la somme des distances intra-nuage :

$$J = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m d^2(x_j, c_i) \quad (3.1)$$

sous la contrainte suivante :

$$\sum_{i=1}^C u_{ij} = 1 \quad \forall j \in [0, N]. \quad (3.2)$$

$m \in [0, \infty]$  caractérise le degré de flou (pour  $m = 0$  on retrouve l'algorithme du *Kmeans*).

$C$  représente le nombre de nuages de points.

$N$  le nombre de pixels.

$c_i$  est le vecteur caractéristique du centre de gravité du nuage de points  $i$ .

$d^2(x_j, c_i)$  est la distance entre  $x_j$  et le centre de gravité du nuage de points  $i$ . Nous utilisons la distance euclidienne.

$C$  et  $m$  sont fixés par l'opérateur une fois pour toute. La minimisation de  $J$  se déroule en deux étapes. Dans la première étape, nous minimisons par rapport aux  $c_i$ . Puis, les  $c_i$  étant fixés, nous minimisons par rapport aux  $u_{ij}$ .

L'algorithme est présenté ci-dessous :

### Algorithme 3.1.1 (Algorithme du *fuzzy Cmeans*)

- ① Initialiser les centres de gravité des nuages de points,  $n = 0$ .
- ② Estimer la matrice de partition  $\mathbf{U}^{(n)}$ .
- ③ Estimer les centres de gravité  $c_i$  des nuages de points.
- ④ Si  $\|\mathbf{U}^{(n)} - \mathbf{U}^{(n-1)}\| > \xi$  alors  $n = n + 1$  et retourner à l'étape 2.  
Sinon aller à l'étape 5.
- ⑤ Assigner à chaque pixel l'étiquette du nuage de points auquel il a la plus grande probabilité d'appartenir.

Les éléments de la matrice de partition sont mis à jour de la manière suivante :

$$\left\{ \begin{array}{l} u_{ij} = \left( \sum_{k=1}^C \left( \frac{d^2(x_j, c_i)}{d^2(x_j, c_k)} \right)^{\frac{1}{m-1}} \right)^{-1} \quad si \quad I_j = \emptyset \\ \left\{ \begin{array}{l} u_{ij} = 0 \quad i \notin I_j \\ u_{ij} = \frac{1}{|I_j|} \quad i \in I_j \end{array} \right. \quad si \quad I_j \neq \emptyset \end{array} \right. \quad (3.3)$$

avec  $I_j = \{i | 1 \leq i \leq C, d^2(x_j, c_i) = 0\}$

### 3.1.3 Critère entropique

Dans nombre de problèmes de partition, le choix du nombre de nuages de points est difficile. En effet, que ce soit en imagerie satellitale ou en imagerie médicale pour un problème donné le nombre de nuages de points “idéal” peut varier. Ainsi, en télédétection, si l’on cherche à distinguer le bâti du non bâti dans certain cas le nombre de nuages de points optimal est supérieur à 2. On peut avoir plusieurs nuages de points caractérisant le bâti ou/et le non bâti. De même en médical, pour la partition du cerveau le nombre de nuages de points se situe aux alentours de 5 (parfois 4, parfois 7). Il paraît donc nécessaire de définir un critère afin de déterminer un nombre de nuages de points de manière optimale et non de manière a priori.

Les approches classiques optimisent un critère de validité. Ce critère qui tient compte des caractéristiques géométriques et topologiques des nuages de points résultants permet d’évaluer la valeur de la classification. Le plus connu est le critère AIC (*A-Information Criterion*) proposé par Akaike dans [2] dans le but d’identifier des modèles statistiques. Ce critère a été repris puis modifié dans [16] afin d’estimer le nombre de textures et les paramètres associés.

Dans [58, 93] les auteurs utilisent un critère qui tient compte de la compacité des nuages de points et de la distance entre eux.

D’autres approches consistent à minimiser un critère. Un des plus connus est le critère MDL (*Maximum Description Length*) proposé par Rissanen [102, 103]. Ce critère a été utilisé en imagerie médicale pour estimer le nombre de classes d’un problème de segmentation de tissus [80]. Cette liste n’est pas exhaustive et le lecteur pourra se référer à l’article de Olivier et al. [94] qui propose une étude comparative de divers critères de type log-vraisemblance pénalisée permettant d’estimer l’ordre d’un modèle. Les auteurs soulignent notamment la tendance à la surparamétrisation du critère AIC.

Frigui et Krishnapuram [39] ont utilisé des concepts issus des statistiques robustes. Leur critère est la somme de 2 termes. Un terme contraint les degrés d'appartenance pour introduire de la "compétition" entre les nuages de points. L'autre terme est défini de manière à prendre en compte les avantages d'une partition hiérarchique.

Palubinskas et al. [95] ont également travaillé sur le sujet dans le cas non flou. Le premier terme de leur critère est la somme des distances intra-nuage. Le second terme est l'entropie de l'histogramme de l'image partitionnée. En restauration d'image, l'entropie a été utilisée comme terme a priori dans le cadre bayésien [51] afin de régulariser l'image. Ces méthodes sont connues sous le nom de méthodes du maximum d'entropie et sont basées sur la théorie de Shannon [66, 113]. En effet, en régularisation le but est de maximiser l'entropie de l'image.

Notre objectif est de minimiser le nombre de nuages de points et donc de minimiser l'entropie de l'histogramme de l'image partitionnée. C'est pourquoi, nous définissons le terme a priori comme suit :

$$\sum_{i=1}^C p_i \log(p_i) , \quad (3.4)$$

où  $p_i$  est la probabilité a priori du nuage de points  $i$ .

Notons  $\mathbf{p}_j = \frac{1}{N}$  la probabilité a priori de chaque pixel  $j$ .

Dans ce qui suit et pour simplifier les notations nous noterons  $j \in i$  pour signifier que le pixel  $j$  appartient au nuage de points  $i$ . Le théorème des probabilités totales [92] conduit aux équations suivantes :

$$\begin{aligned} p_i &= \sum_{j=1}^N \mathbf{p}_j p(j \in i) \\ p_i &= \sum_{j=1}^N \frac{1}{N} p(j \in i) \end{aligned} \quad (3.5)$$

Dans [104], Ruspini interprète  $u_{ij}$  comme la probabilité  $p(j \in i)$  que l'objet  $j$  appartienne au nuage de points  $i$ . En effet, il souligne le fait que dans le cas flou le degré d'appartenance a une forte signification probabiliste. Il met également en évidence l'utilité, en reconnaissance de formes, du passage de l'une des représentations (degré d'appartenance) à l'autre (probabilité).

De même, dans ce travail, nous interprétons  $u_{ij}$  comme la probabilité  $p(j \in i)$  que le pixel  $j$  appartienne au nuage de points  $i$ .

Notons que :  $u_{ij} \in [0, 1]$  et  $\sum_{i=1}^C u_{ij} = 1, \quad \forall j$ .

La probabilité a priori du nuage de points  $i$  est donc définie comme suit :

$$p_i = \frac{1}{N} \sum_{j=1}^N u_{ij}. \quad (3.6)$$

Le critère  $J$  est défini par :

$$J = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m d^2(x_j, c_i) - \alpha \sum_{i=1}^C p_i \log(p_i) \quad (3.7)$$

$J$  est minimisé sous la contrainte suivante :

$$\sum_{i=1}^C u_{ij} = 1 \quad \forall j \in [0, N]. \quad (3.8)$$

Dans la suite du manuscrit, on prend  $m = 2$ . Cette valeur est valable pour la plupart des applications et est très souvent utilisée en classification floue.

$C$  est le nombre initial de nuages de points choisi par l'opérateur ( $C$  peut être le nombre de niveaux de gris). Ce nombre décroît à chaque itération jusqu'à atteindre la valeur optimale  $C_{opt}$ . C'est la principale différence avec l'algorithme classique pour lequel le nombre de nuages de points est fixe.

Le premier terme de  $J$  caractérise l'homogénéité des nuages de points. Ce terme est minimal lorsqu'un nuage de points caractérise un seul niveau de gris.

Le second terme est l'entropie de l'histogramme de l'image partitionnée. Ce terme est minimal lorsque tous les pixels sont dans le même nuage de points et que tous les autres sont vides.

Nous introduisons la contrainte (3.8) dans le critère en utilisant les multiplicateurs de Lagrange. Finalement, le critère que nous minimisons est le suivant :

$$\begin{aligned} J &= \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^2 d^2(x_j, c_i) - \alpha \sum_{i=1}^C p_i \log(p_i) \\ &- \sum_{j=1}^N \lambda_j \left( \sum_{i=1}^C u_{ij} - 1 \right). \end{aligned} \quad (3.9)$$

Nous minimisons  $J$  tour à tour par rapport aux  $c_i$  et par rapport aux  $u_{ij}$ . Nous annulons les dérivées partielles de  $J$  pour obtenir les équations de mise à jour des paramètres  $c_i$  et  $u_{ij}$  :

1. Les degrés d'appartenance  $u_{ij}$  étant connus, on minimise  $J$  par rapport aux centres de gravité  $c_i$  :

$$\frac{\partial J}{\partial c_i} = \sum_{j=1}^N (u_{ij})^2 \frac{\partial d^2(x_j, c_i)}{\partial c_i} = 0 \quad (3.10)$$

2. Les centres de gravité  $c_i$  étant connus, on minimise  $J$  par rapport aux  $u_{ij}$  :

$$\frac{\partial J}{\partial u_{ij}} = 2u_{ij}d^2(x_j, c_i) - \alpha(1 + \log(p_i)) \frac{\partial p_i}{\partial u_{ij}} - \lambda_j = 0 \quad (3.11)$$

$$\frac{\partial p_k}{\partial u_{ij}} = \frac{\delta(k, i)}{N} \quad \text{où } \delta \text{ est le symbole de Kronecker}$$

Les degrés d'appartenance et les probabilités a priori de chaque nuage de points ne varient pas significativement d'une itération à l'autre. Ainsi, les paramètres  $c_i$  et  $u_{ij}$  à l'itération  $n$  sont mis à jour de la manière suivante :

$$c_i^{(n)} = \frac{\sum_{j=1}^N [u_{ij}^{(n-1)}]^2 x_j}{\sum_{j=1}^N [u_{ij}^{(n-1)}]^2}, \quad (3.12)$$

$$u_{ij}^{(n)} = \frac{\frac{1}{d^2(x_j, c_i^{(n)})}}{\sum_{k=1}^C \frac{1}{d^2(x_j, c_k^{(n)})}} + \frac{\alpha \left( 1 + \log(p_i^{(n-1)}) - \frac{\sum_{k=1}^C \frac{1 + \log(p_k^{(n-1)})}{d^2(x_j, c_k^{(n)})}}{\sum_{k=1}^C \frac{1}{d^2(x_j, c_k^{(n)})}} \right)}{2N d^2(x_j, c_i^{(n)})}. \quad (3.13)$$

Il est préférable de commencer l'algorithme avec une grande valeur de  $\alpha$  afin de déterminer dès les premières itérations le nombre optimal de nuages de points puis de tenir compte peu à peu du terme de classification. Pour ce faire,  $\alpha$  décroît de manière exponentielle [39] à chaque itération de la manière suivante :  $\alpha(n) = \alpha(0) \exp\left(-\frac{n}{\tau}\right)$ , où  $n$  est le nombre d'itérations et  $\tau$  est une constante de temps.

On évite ainsi que le terme de classification ne soit biaisé par le terme d'entropie puisque  $\alpha(n)$  tend vers 0 quand  $n$  tend vers  $\infty$ .

En outre,  $\alpha(0)$  dépend fortement du nombre de pixels dans l'image. Afin de s'affranchir des données, nous posons  $\alpha(0) = \alpha_0 N$ . Dans ce cas, le seul paramètre que nous avons à fixer est  $\alpha_0$ . Ce paramètre détermine la finesse de la classification que l'on souhaite obtenir. Pour une application donnée, nous fixons une fois pour toute  $\alpha_0$  et  $\tau$  alors que le nombre de nuages de points peut être différent suivant la scène traitée.

Après avoir mis à jour les centres de gravité et les degrés d'appartenance, nous estimons les probabilités a priori de chacun des nuages de points. Nous éliminons les nuages de points dont la probabilité est inférieure à un certain seuil  $\epsilon$ .

L'algorithme est présenté ci-dessous :

### Algorithme 3.1.2 (Algorithme du *fuzzy Cmeans* modifié)

- ① Initialiser les centres de gravité des nuages de points,  $n = 0$ .
- ② Calculer la matrice de partition  $\mathbf{U}^{(n)}$  (cf. équation 3.13).
- ③ Calculer les centres de gravité des nuages de points (cf. équation 3.12).
- ④ Calculer les probabilités a priori  $p_i$  de chaque nuage de points (cf. équation 3.6),  
si  $p_i < \epsilon$  le nuage de points  $i$  est supprimé.
- ⑤ Si  $\|\mathbf{U}^{(n)} - \mathbf{U}^{(n-1)}\| > \xi$  alors  $n = n + 1$ , mettre à jour  $\alpha$  et retourner à l'étape 2.  
Sinon aller à l'étape 6.
- ⑥ Assigner à chaque pixel l'étiquette du nuage de points auquel il a la plus grande probabilité d'appartenir.

### 3.1.4 Résultats expérimentaux

Afin de faciliter les commentaires, nous noterons KM l'algorithme du *Kmeans* (*Fuzzy Cmeans* version "dure"), FCM l'algorithme du *Fuzzy Cmeans* classique, et FCME l'algorithme du *Fuzzy Cmeans* avec terme entropique.

La méthode consiste donc à initialiser  $C$  avec un nombre élevé (par exemple le nombre de niveaux de gris dans l'image) et à le faire décroître à chaque itération jusqu'à atteindre le nombre de nuages de points optimal  $C_{opt}$ . Pour des questions pratiques (temps de calcul notamment),  $C$  est inférieur au nombre de niveaux de gris. Ce choix peut être guidé par l'application si l'on a une vague idée du nombre de nuages de points recherchés.

Dans les tests qui suivent nous avons fait tourner l'algorithme FCM pendant 2 itérations avec le nombre de nuages de points initial afin d'initialiser l'algorithme modifié avec une classification initiale correcte.

Le premier exemple (Figure 3.1 (a)) est la partition d'une image SPOT3 de champs. Nous voulons partitionner les données mais nous ne connaissons pas a priori le nombre de nuages de points. Nous choisissons d'initialiser l'algorithme avec un grand nombre de nuages de points (40, ou 50). L'algorithme FCME détermine le nombre optimal de nuages de points ( $C_{opt} = 9$ ) dès les premières itérations (Figure 3.4 (b)). Soulignons que le résultat obtenu ne dépend pas du nombre initial choisi pour  $C$  (40 ou 50).

Le résultat de la partition par l'algorithme FCME est une matrice de partition. Nous disposons donc de 9 images pour l'exemple ci-dessus (une par nuage de points), pour lesquelles la valeur associée à un pixel est le degré d'appartenance de ce pixel au nuage de points en question. Il est cependant peu informatif de toutes les présenter. Figure 3.2, nous présentons 3 exemples (3 nuages de points parmi 9). Par ailleurs, l'ultime étape de l'algorithme consiste à associer à chaque pixel l'étiquette du nuage de points auquel il a la plus grande chance d'appartenir. Dans ce cas, pour l'image 3.1 (a) le résultat est donné figure 3.4 (b).

Notons que le résultat de la classification finale ne diffère pas de celui qui est obtenu par l'algorithme FCM pour lequel le nombre de nuages de points est supposé connu et égal à 9 car le paramètre  $\alpha$  décroît avec le nombre d'itérations.

La figure 3.4 (a) représente le résultat de la classification obtenue par l'algorithme KM avec un nombre de nuages de points égal à 9. Dans ce cas, la classification diffère de celle obtenue par FCME (Figure 3.4 (b)). Elle est plus bruitée. La différence provient non pas du terme d'entropie rajouté<sup>1</sup> mais de l'aspect flou. En

---

<sup>1</sup>il n'y a pas de différence entre la classification obtenue par FCM et par FCME.

effet, avec FCM ou FCME la décision de la classification constitue l'ultime étape de l'algorithme. Enfin, notons que pour cet exemple et pour le suivant le nombre de nuages de points peut être difficile à déterminer, voire impossible lorsque que l'on traite des problèmes dont on ignore le nombre de nuages de points. Dans ce cas il est très intéressant d'utiliser l'algorithme FCME.

En terme de coût de calcul, l'algorithme FCME est plus demandeur dans la mesure où nous l'initialisons avec la matrice de partition obtenue après avoir fait tourné l'algorithme FCM avec le nombre initial de nuages de points  $C$  qui peut être grand (40 ou 50 dans l'exemple ci-dessus). En revanche, le corps même de l'algorithme FCME ne diffère d'avec celui de l'algorithme du FCM que par le calcul des probabilités a priori de chaque nuage de points, ce qui ne constitue pas un calcul conséquent.

Afin de montrer le large champ d'application de l'algorithme proposé, nous l'avons appliqué à des données IRM<sup>2</sup> du cerveau. L'image initiale est présentée figure 3.1 (b). Le nombre initial de nuages de points a été fixé à 10, 12, et 14. Les résultats (Figure 3.4 (d)) sont les mêmes quelque soit l'initialisation. Le nombre optimal de nuages de points déterminé par l'algorithme est  $C_{opt} = 6$ . La matière blanche et la matière grise apparaissent clairement sur la figure 3.4 (d). Nous pouvons faire exactement les mêmes remarques que pour l'exemple précédent. Dans ce cas précis, il n'est pas évident de déterminer a priori le nombre de nuages de points, notamment à cause des différents tissus et graisses entourant le crâne.

Le dernier exemple est un problème d'extraction de zones urbaines. Nous estimons le paramètre de texture qui est caractéristique des zones urbaines. La figure 3.5 (a) est une région dans laquelle il y a une zone urbaine. La figure 3.5 (b) est une région dans laquelle il n'y a pas de zone urbaine. Les figures 3.5 (c) et (d) correspondent aux estimées du paramètre de texture. Nous appliquons les algorithmes KM et FCME avec un nombre initial de nuages de points égal à 2.

Nous faisons tourner l'algorithme KM. Nous obtenons les partitions représentées par les figures 3.7 (a) et (b). 2 nuages de points sont détectés qu'il y ait une ville ou non car ce nombre est fixé a priori.

Avec l'algorithme FCME, un seul nuage de points, est détecté lorsqu'il n'y a pas de zone urbaine (Figure 3.7 (d)) et deux dans le cas contraire (Figure 3.7 (c)). Les

---

<sup>2</sup>Nous remercions le Dr. Dormont de La Pitié Salpêtrière pour les données.

$\alpha_0$	$\tau$	$\epsilon$	$\xi$
2.	10	$10^{-3}$	$10^{-5}$

**Tableau 3.1:** Paramètres de l'algorithme FCME.

figures 3.5 (e) et (f) représentent les histogrammes des images du paramètre de texture estimé. Sur la figure (f) nous avons bien un seul mode. En revanche, sur la figure (e) nous avons un mode très marqué et un second mode qui l'est moins.

Dans l'exemple (Figure 3.5 (b)) ci-dessus, nous sommes dans un cas où nous surestimons le nombre de nuages de points. De même, nous avons toujours le risque de sous-estimer le nombre de nuages de points pour un problème dont on ne connaît pas le nombre exact de nuages de points. Il est donc intéressant d'avoir l'algorithme FCME pour éviter ce genre de problèmes.

En outre, en ce qui concerne l'application traitée, il peut arriver que l'histogramme du paramètre de texture estimé ne soit pas bimodal. Dans ce cas, si l'on applique KM avec un nombre de nuages de points égal à 2, la classification obtenue est rarement satisfaisante. Dans ce cas, une classification avec un nombre de nuages  $C_{opt}$  supérieur à 2, suivi d'un regroupement des nuages de points de manière supervisée de manière à obtenir les 2 classes : classe urbaine et classe non urbaine, est nécessaire. Le nombre  $C_{opt}$  peut alors varier d'une image à l'autre. L'algorithme FCME permet de l'estimer automatiquement.

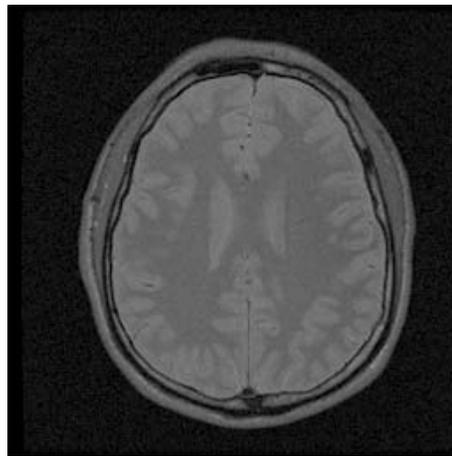
Finalement, notre objectif étant de segmenter une image satellitale afin d'extraire un masque urbain, le résultat qui nous intéresse pour la suite de notre travail est l'estimation du nombre de nuages de points et la matrice de partition. En effet, ces paramètres vont être intégrés dans un modèle de segmentation présenté dans le paragraphe 3.2. La classification obtenue à cette étape va nous permettre d'initialiser l'algorithme de segmentation.

## 3.2 Régularisation par modélisation markovienne

Nous utilisons la modélisation markovienne pour introduire de l'information contextuelle dans le processus de classification et donc régulariser l'image précédente. Comme le souligne Gouinaud dans [47], "l'utilisation des champs markoviens pour la classification urbaine permet d'améliorer les résultats [...]. L'amélioration du résultat pour la détection du masque urbain est surtout perceptible au niveau de la cohérence spatiale".

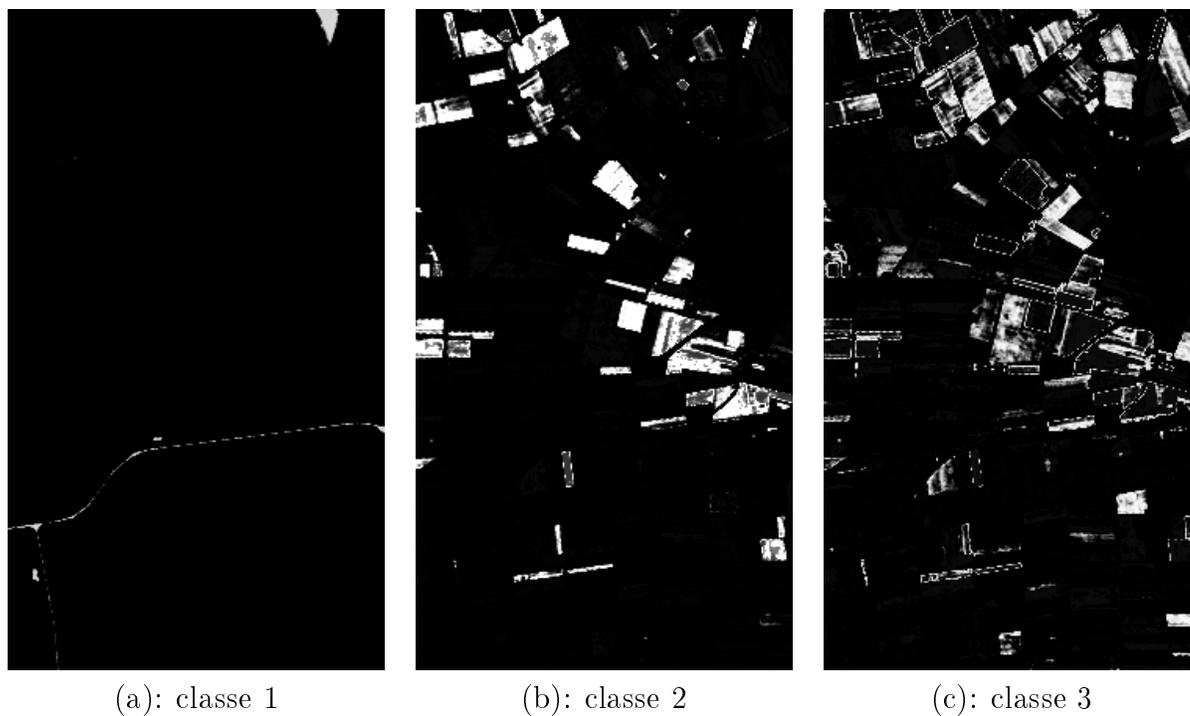


(a)

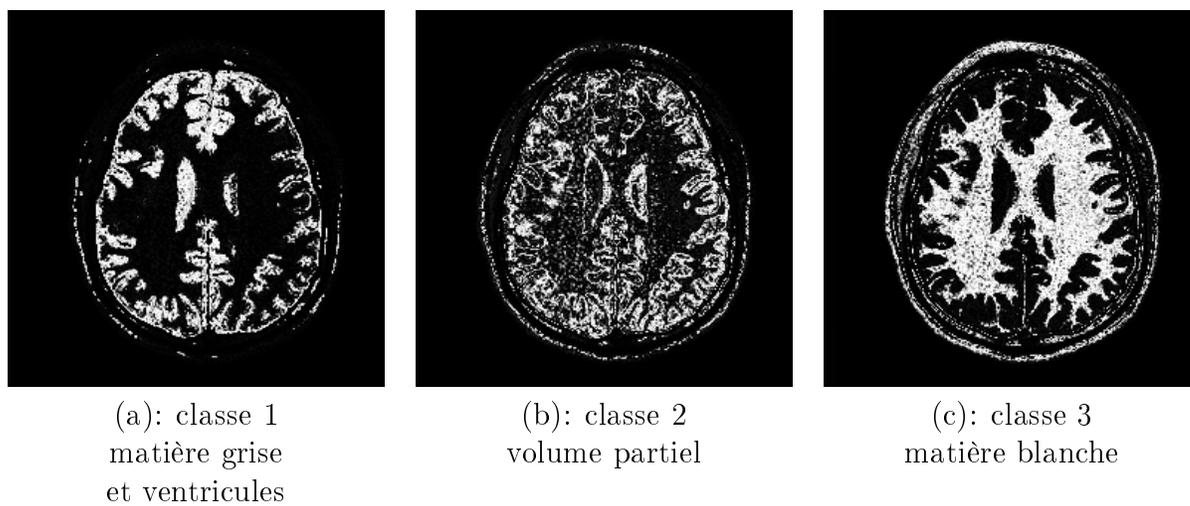


(b)

**Figure 3.1:** (a): Image panchromatique SPOT3 (b): MRI d'un cerveau (echo T1 d'un scanner sigma 1.5 Tesla, GE).



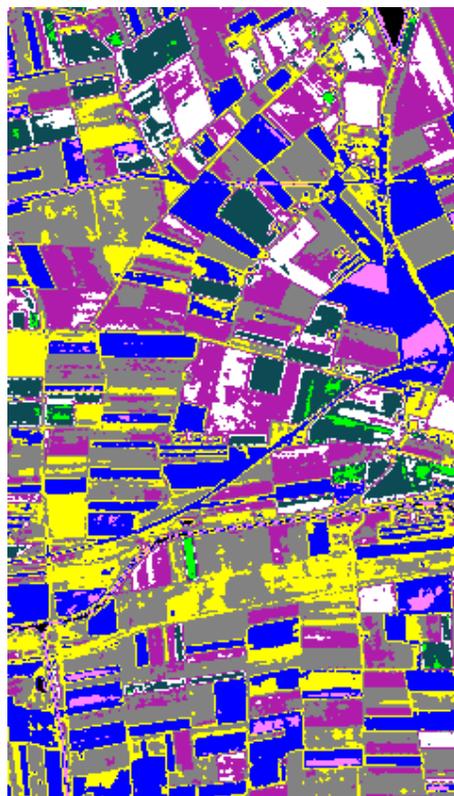
**Figure 3.2:** Exemples d'images des degrés d'appartenance à 3 classes : image SPOT3.



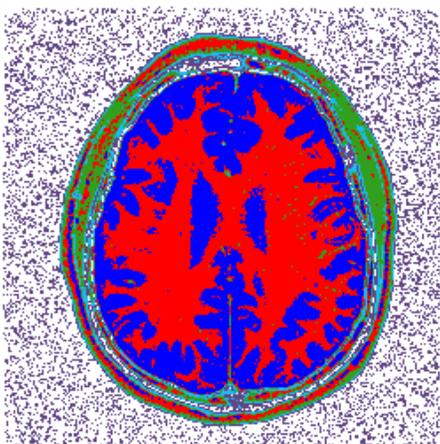
**Figure 3.3:** Exemples d'images des degrés d'appartenance à 3 classes : IRM d'un cerveau.



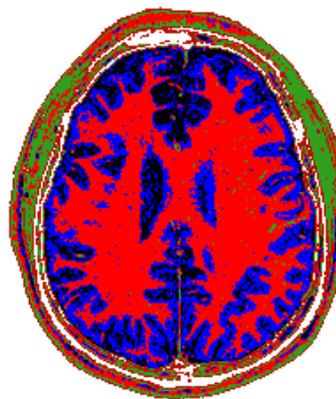
(a):  $C = 9$ ,  $CPU = 3.53s$



(b):  $C_{max} = 40$ ,  $C_{opt} = 9$   
 $CPU = 325.86s$



(c)  $C = 6$ ,  $CPU = 3.14s$



(d):  $C_{max} = 10$ ,  $C_{opt} = 6$   
 $CPU = 47.80s$

**Figure 3.4:** Classification (a),(c): avec KM (b),(d): avec FCME.



(a): Zones urbaine



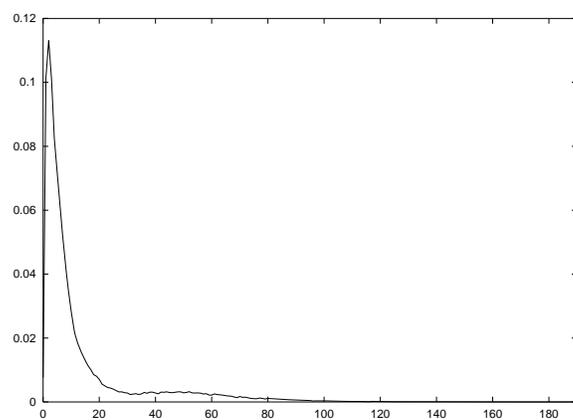
(b): Pas de zone urbaine



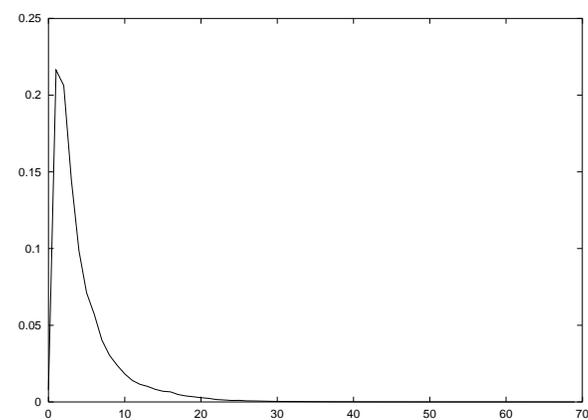
(c)



(d)



(e)



(f)

**Figure 3.5:** (a), (b): Simulations SPOT5. (c) et (d): Images du paramètre de texture estimé. (e) et (f) histogrammes des images (c) et (d).

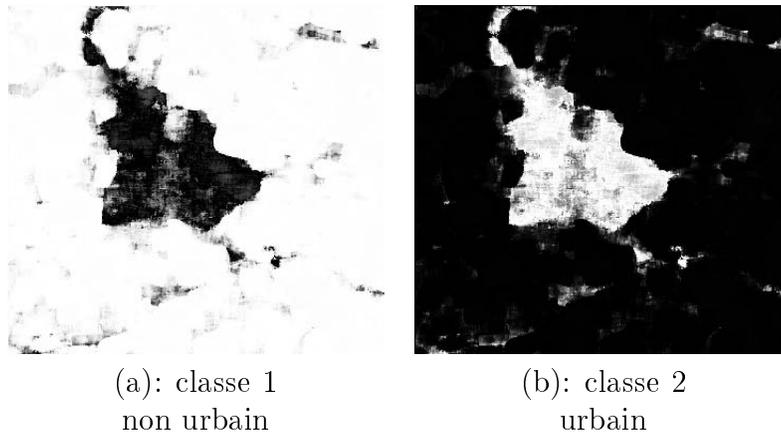
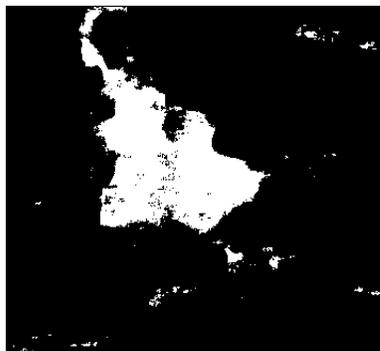


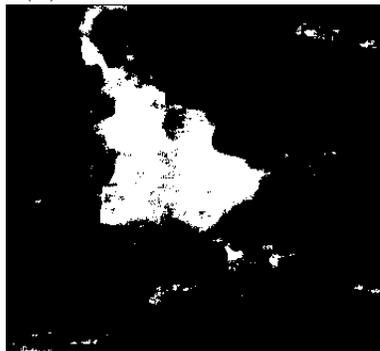
Figure 3.6: Images des degrés d'appartenance à chacune des 2 classes (FCME).



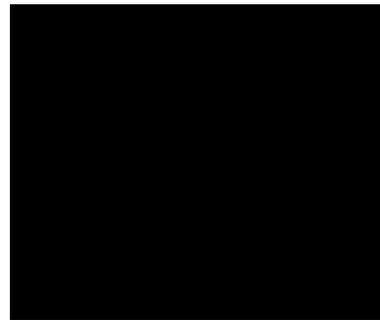
(a):  $C = 2$ , CPU = 1.56s



(b):  $C = 2$ , CPU = 3.12s



(c):  $C_{opt} = 2$ ,  $C_{max} = 12$ , CPU = 78.18s



(d):  $C_{opt} = 1$ ,  $C_{max} = 12$ , CPU = 60.86s

Figure 3.7: Classification (a),(b): avec KM (c),(d): avec FCME.

### 3.2.1 Modèle de segmentation

#### 3.2.1.1 Critère du MAP

Après avoir défini les champs de Markov et les distributions de Gibbs (paragraphe 1.2.2.4), nous considérons le contexte particulier de la segmentation d'image.

Notons  $\Omega$  l'ensemble des états des observations et  $\Lambda$  l'ensemble des étiquettes. Les états font référence ici aux valeurs possibles des paramètres de texture.

Nous notons  $X$  le champ aléatoire relatif à l'image observée à valeurs dans  $\Omega^N$ .

Nous notons  $L$  le champ aléatoire relatif à l'image étiquetée à valeurs dans  $\Lambda^N$ .

$X$  étant connu, l'estimateur du MAP est la configuration  $\hat{L}$  qui maximise l'expression suivante :

$$P(L|X) = \frac{P(X|L)P(L)}{P(X)}$$

ce qui revient,  $X$  étant connu et donc  $P(X)$  constant, à maximiser l'expression suivante :

$$P(X|L)P(L)$$

Dans la suite, nous définissons successivement la vraisemblance  $P(X|L)$  et le terme a priori  $P(L)$ . En théorie et sous certaines conditions de décroissance en température, la configuration optimale  $\hat{L}$  qui maximise  $P(X|L)P(L)$  peut être obtenue par l'algorithme du recuit simulé [43]. Cependant, pour des questions de coût de calcul, nous utiliserons un algorithme déterministe, l'ICM (*Iterated conditional modes*), décrit dans le paragraphe 3.2.2, qui permet d'obtenir un minimum local mais qui donne des résultats tout à fait satisfaisants si l'algorithme est initialisé par une image proche de la solution cherchée.

#### 3.2.1.2 Vraisemblance

Dans ce paragraphe, nous comparons deux approches pour modéliser  $P(X|L)$ . Pour cela nous utilisons l'information obtenue par l'algorithme du *fuzzy Cmeans* modifié.

### a. Premier modèle

Dans cette approche, nous considérons la fonction d'appartenance de la partition floue comme représentant la vraisemblance. La loi d'attache aux données considérée est la suivante :

$$P(X|L) = \prod_{s \in S} \left( \sum_{j \in \Lambda} (u_{js} \delta(l_s, j)) \right) \quad (3.14)$$

$u_{js}$  est le degré d'appartenance du pixel  $s$  à la classe  $j$ .

$l_s$  est l'étiquette associé au pixel  $s$ .

$$\delta(l_s, j) = \begin{cases} 1 & , \text{ si } l_s = j \\ 0 & , \text{ si } l_s \neq j \end{cases}$$

$u_{js}$  est obtenu par l'algorithme FCME.

### b. Second modèle

Nous avons voulu comparer les résultats obtenus avec le premier modèle avec ceux obtenus en modifiant le terme d'attache aux données. Dans ce second modèle, nous considérons une attache aux données gaussienne dont les paramètres (moyenne et variance des classes) sont estimés à partir de la partition obtenue par l'algorithme du *fuzzy Cmeans*. La loi  $P(X|L)$  est définie comme suit :

$$P(X|L) = \prod_{s \in S} \left( \sum_{j \in \Lambda} \left( \frac{\delta(l_s, j)}{\sqrt{2\pi}\sigma_j} \exp \left( -\frac{(x_s - \mu_j)^2}{2\sigma_j^2} \right) \right) \right) , \quad (3.15)$$

$l_s$  est l'étiquette associé au pixel  $s$ .

$$\delta(l_s, j) = \begin{cases} 1 & , \text{ si } l_s = j \\ 0 & , \text{ si } l_s \neq j \end{cases}$$

$\sigma_j^2$  et  $\mu_j$  sont respectivement la variance et la moyenne de la classe  $j$  et sont estimés par les estimateurs empiriques suivants :

$$\mu_j = \frac{\sum_{s=1}^N u_{js}^2 x_s}{\sum_{s=1}^N u_{js}^2}$$

$$\sigma_j^2 = \frac{\sum_{s=1}^N u_{js}^2 (x_s - \mu_j)^2}{\sum_{s=1}^N u_{js}^2}$$

### 3.2.1.3 A priori

Le modèle a priori considéré est le modèle de Potts [43, 114]. La probabilité a priori est la suivante :

$$P(L) = \frac{1}{Z_\beta} \exp \left( \beta \sum_{c=\{s,t\}, c \in \mathcal{C}} \delta(l_s, l_t) \right), \quad (3.16)$$

$$\delta(l_s, l_t) = \begin{cases} 1 & , \text{ si } l_s = l_t \\ 0 & , \text{ si } l_s \neq l_t \end{cases}$$

Dans ce qui suit le paramètre  $\beta$  est fixé par essais successifs. A noter qu'il est choisi une fois pour toute pour tous les tests expérimentaux. Les seules cliques dont nous tenons compte sont les paires de sites en 8-connexité.

## 3.2.2 Algorithme ICM (*Iterated conditional modes*)

L'algorithme ICM [13], également appelé échantillonnage à température nulle [128], consiste à affecter au pixel  $s$  l'étiquette de la classe qui conduit à maximiser la probabilité conditionnelle en  $s$ , i.e.  $P(X_s | L_s)P(L_s | L_r, r \in V(s))$ .

Le résultat obtenu dépend fortement de l'initialisation et du choix de l'ordre de balayage de l'image. C'est pourquoi dans les expériences qui suivent, l'algorithme

sera initialisé avec le résultat de la classification obtenue suite à l'algorithme FCME. L'avantage de l'ICM est son faible coût en temps de calcul comparé aux autres algorithmes combinatoires (e.g. recuit simulé, échantillonneur de Gibbs).

Dans [37], Forbes souligne que si l'on dispose d'une segmentation initiale (ou que l'image originale est discrète) et si l'on fait l'hypothèse du modèle de Potts et d'un modèle bruit, alors elle montre que :

- L'algorithme ICM est équivalent à une forme de filtrage morphologique.
- La segmentation est peu sensible à des valeurs précises des paramètres du modèle.

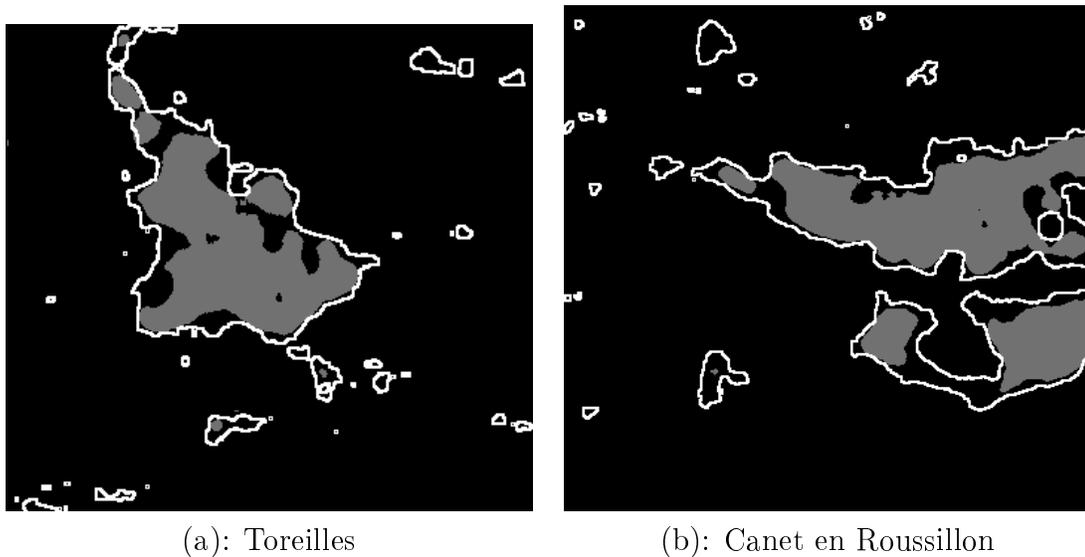
Nous avons pu le vérifier. En effet, lorsque nous avons appliqué l'algorithme ICM les résultats étaient peu sensibles au choix de  $\beta$ . Nous avons également testé l'algorithme du recuit simulé avec la dynamique de Métropolis. Dans ce cas, le choix de  $\beta$  a été primordial. Par ailleurs, pour des résultats tout aussi satisfaisants, l'algorithme déterministe ICM est bien moins long (quelques secondes) contre de nombreuses minutes pour l'algorithme combinatoire.

### 3.2.3 Image de marqueurs

Dans le chapitre 2.4.4 nous avons précisé que nous gardions l'image pour laquelle le niveau de gris de chaque pixel est la valeur minimale des 8 variances conditionnelles estimées. Cette image est partitionnée par l'algorithme FCME. Nous filtrons ensuite l'image partitionnée avec un filtre médian. Le résultat nous donne les limites intérieures de la ville. Cette image nous sert à marquer la présence d'une zone urbaine. Nous ne conservons donc de la segmentation précédente que les régions marquées (Figures 3.8 (a) et (b)), i.e. les zones détectées comme ville et qui ont une intersection non nulle avec le résultat du filtre médian. Cette étape nous permet donc d'éliminer la plupart des fausses alarmes.

## 3.3 Résultats

L'algorithme global, appelé EXZUM (EXtraction des Zones Urbaines par modélisation Markovienne), est décrit ci-dessous :



**Figure 3.8:** Segmentation obtenue par l'ICM (contours blancs) et marqueurs (régions grises).

#### Algorithme 3.3.1 (Algorithme EXZUM)

- ① Estimer les variances conditionnelles locales dans les huit directions.
- ② Classer ces 8 valeurs dans l'ordre croissant.
- ③ Conserver une combinaison linéaire des 2 valeurs centrales.
- ④ Appliquer l'algorithme du fuzzy Cmeans modifié.
- ⑤ Régulariser par champs de Markov.
- ⑥ Éliminer les fausses alarmes avec l'image des marqueurs.

Des tests ont été effectués sur des simulations SPOT5 (Figures 3.11, 3.12, et 3.14), fournies par le CNES. La résolution des images utilisées est de 5 mètres exceptée l'image de Toulouse (Figure 3.14) qui est à une résolution de 2m50. Il s'agit dans tous les cas du canal panchromatique. Nous avons également fait des tests sur des images SPOT3 (Figures 3.15 et 3.16) à 10 mètres de résolution et sur des images radar ERS1 à 12,5 mètres de résolution (Figures 3.17 et 3.19).

Dans tous les tests qui suivent la taille de la fenêtre pour l'estimation des paramètres de texture dépend de l'image. Cette taille a été déterminée par essais successifs. A noter qu'elle dépend de la résolution de l'image, du type de capteur mais également de la taille des structures de la ville. Tous les paramètres ainsi que les temps de calcul de chaque étape, et ce pour chaque image, sont précisés dans le tableau 3.2.

Les résultats obtenus à partir des images optiques (SPOT3 et SPOT5) sont satisfaisants dans la mesure où les limites de la ville sont bien localisées et que le nombre de fausses alarmes est faible. La forme géométrique des zones de bâti est bien respectée. Même les zones de forme plus longiligne sont bien détectées (Figure 3.15).

Dans le cas des simulations SPOT5, la seule validation que nous avons pu effectuer était de comparer d'un point de vue qualitatif les résultats obtenus et les images aériennes multi-canal à notre disposition. Cela nous a permis de vérifier que le contour des villes est correct, que la géométrie est respectée et que les zones de bâtis détectées sont effectivement des zones de bâtis. En ce qui concerne les images SPOT3, nous n'avions d'autre moyen que de comparer visuellement les contours détectés automatiquement avec l'image satellitale elle-même.

Par ailleurs, nous obtenons de meilleurs résultats en utilisant le second terme d'attache aux données. En effet, les contours sont plus précis (Figures 3.9 (a),(b) et 3.10 (a)(b)). Cependant, il reste sur ces images des fausses alarmes. Nous utilisons donc une image de marqueurs (Figures 3.8 (a) et (b)) afin de les éliminer.

Au départ notre objectif était d'extraire un masque urbain sur des images optiques. La méthode développée permet également (Figures 3.11,3.14) de distinguer dans l'urbain toutes les zones végétales de surfaces suffisamment grandes, qui présentent les mêmes caractéristiques que les champs, i.e. zones de niveaux de gris uniformes.

Des tests ont été effectués sur l'image du canal XS3 de l'image aérienne de Toreilles à 1m67 de résolution. Sur cette image les résultats ne sont pas bons. Des erreurs importantes sont commises sur les frontières de la zone urbaine. Ceci n'a rien de surprenant dans la mesure où, à cette résolution, la notion même de texture au sens de la répétition périodique d'un même motif n'a pas lieu d'être. Il nous paraît nécessaire d'utiliser des méthodes basées sur des critères géométriques (ex : géométrie probabiliste ...) afin d'extraire directement les bâtiments.

Par ailleurs, grâce à l'aide de l'équipe du Pr. Rudant de l'université de

Marne-La-Vallée (UMLV), nous avons pu obtenir des images radar ERS1 des villes de Belém et Macapá situées au Brésil (Figures 3.17 et 3.19). Comme beaucoup d'algorithmes de traitement d'image fonctionnent sur des images codées sur 8 bits, les utilisateurs sont souvent amenés à compresser leurs images sur 8 bits. Dans le cas présent, la méthode utilisée a consisté à ramener les valeurs des niveaux de gris entre 0 et 255 en divisant les valeurs par 3 et en saturant les valeurs supérieures à 255. Les images ont également été filtrées par le filtre de speckle de Crimmins [23]. Notons que le passage d'un codage sur 16 bits à un codage sur 8 bits entraîne une perte d'une partie de l'information. Notamment la saturation à 255 n'est pas sans effet sur les statistiques des classes les plus brillantes. Nous pouvons vérifier cela sur certaines régions brillantes de la zone urbaine qui deviennent uniformes (Figures 3.17 (b): extrémité sud-ouest de la ville).

Cependant soulignons également que l'heuristique de l'algorithme d'estimation des variances conditionnelles locales qui nous permet d'obtenir des bords nets requiert un nombre suffisant d'échantillons. Or avec des niveaux de gris variant de 0 à  $2^{16} - 1$ , le nombre d'échantillons pour chaque distribution conditionnelle  $P(X_s|m_s^d)$  sera faible et donc l'estimation de la variance conditionnelle locale peu fiable.

L'algorithme a donc été appliqué dans un premier temps sur des images codées sur 8 bits (Figures 3.17 et 3.19) en tenant compte de l'heuristique puis sur des images codées sur 16 bits (Figures 3.20 et 3.21) en prenant en compte tous les échantillons. Les deux images codées sur 16 bits nous ont été fournies par le groupe Traitement et Interprétation des images du département TSI de l'ENST dans le cadre du PNTS-ATP 98.

Les résultats obtenus avec l'algorithme EXZUM sur des images codées sur 8 bits sont de bonne qualité d'après les thématiciens. Dans les zones d'Amérique du sud cela est d'un grand intérêt dans la mesure où les perturbations atmosphériques posent d'importants problèmes pour l'analyse des données satellitales optiques (de type SPOT). Dans les pays émergents l'intérêt essentiel de cette analyse est la gestion des constructions anarchiques (e.g. bidonville), de la pression urbaine sur le milieu naturel [9].

Les résultats ont été analysés par le Dr. Letourneau, expert géographe de l'ENS Ulm et de UMLV, qui a une bonne connaissance du terrain. La ville de Macapá est ainsi très bien détectée (Figure 3.17). Dans le cas de cette ville, nous avons segmenté deux images prises à deux dates différentes 1992 et 1996 [76]. Ainsi on constate que les quartiers récents ne sont pas retenus comme zone urbaine. En effet, dans ces quartiers, la texture urbaine n'est pas complètement formée, i.e. il reste encore des zones non bâties.

En outre, sur l'imagerie radar certaines fausses alarmes peuvent subsister dans

les zones où la rétro-diffusion est forte et spatialement inhomogène (e.g. dans les zones de chenaux de marées sur l'image prise en 1992 en bas à droite). De même, certains domaines urbains de faible rétro-diffusion (de par l'orientation des toits par exemple) sont moins bien détectés.

Le résultat (Figure 3.20) obtenu sur l'image de Macapá codée sur 16 bits est de meilleure qualité. En effet, sur l'image codée sur 8 bits, les zones saturées à 255 sont constantes et répondent donc faiblement au paramètre de texture. Les résultats obtenus sur la ville d'Aix en Provence (Figure 3.21 image codée sur 16 bits) sont également de bonne qualité, i.e. les zones bâties sont bien détectées.

De manière générale, il faut préciser que la détection est de meilleure qualité (moins de fausses alarmes et contours plus précis) en imagerie optique à condition, bien entendu, qu'il n'y ait pas de nuages. C'est pourquoi il est intéressant que l'analyse automatique fonctionne également en imagerie radar pour traiter des zones souvent recouvertes d'une couche nuageuse comme cela est souvent le cas dans les zones tropicales. Notons que dans le cas de l'imagerie radar l'étude a été réalisée à partir d'une seule image radar. En mai 2000, l'ESA (*European Space Agency*), prévoit de lancer le satellite ENVISAT1 avec, à son bord, plusieurs instruments embarqués qui assureront à la fois une continuité avec ERS ainsi que certaines améliorations. Ainsi, l'instrument ASAR (*Advanced Synthetic Aperture Radar*), développé par l'ESA, aura des capacités supérieures à celles de l'instrument SAR, supporté par ERS, notamment en terme d'étendue des angles d'incidences. Il sera alors très intéressant de voir l'apport d'une analyse multi-incidence pour l'analyse urbaine.

### 3.4 Conclusion

Nous avons donc proposé une méthode d'extraction des zones urbaines fondée sur une analyse de la texture de l'image. Pour ce faire, nous utilisons une approche multi-directionnelle. Ceci est très utile pour notre application car très souvent les petits parcellaires, les hangars ou encore les serres sont confondus avec la ville. Or c'est justement une orientation privilégiée qui les différencie.

L'autre intérêt de la modélisation considérée est que nous manipulons des modèles qui sont très simples et demandent peu de temps de calcul. Le seul paramètre que nous avons à déterminer pour estimer le paramètre de texture est la taille de la fenêtre glissante. Notons que dans [25] Descombes avait introduit un paramètre d'ordre afin d'éliminer les fausses alarmes mais que l'estimation de

celui-ci se traduisait par un accroissement important de la charge de calcul.

Le paramètre proposé est plus efficace que les méthodes classiques d'analyse de texture pour la discrimination bâti/non bâti. Signalons également qu'en plus de permettre de délimiter le contour de la ville, la méthode développée permet de caractériser dans la ville les zones végétales (parcs, pelouse de stade ...).

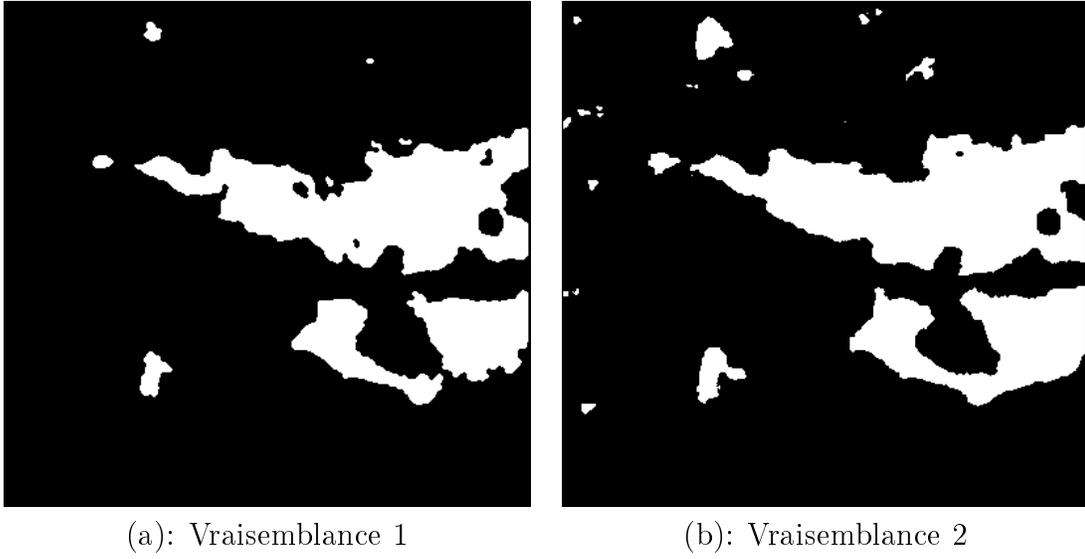
Remarquons, par ailleurs, que le domaine de validité de la méthode proposée semble plus liée à la résolution de l'image traitée (quelques mètres) qu'à la physique du capteur (optique ou radar). Des tests effectués sur des images à plus haute résolution (images aériennes à 1m67) ont montré les limites de notre approche aux résolutions métriques. En effet, à ces résolutions, l'aspect de texture est perdu du fait de la taille des objets (bâtiments, routes ...).

Nous avons également apporté une modification à l'algorithme FCM ce qui permet de ne pas avoir à fixer un nombre de classes précis et surtout de ne pas détecter plusieurs classes lorsqu'il n'y en a qu'une. On élimine ainsi le risque de sous-estimer ce nombre.

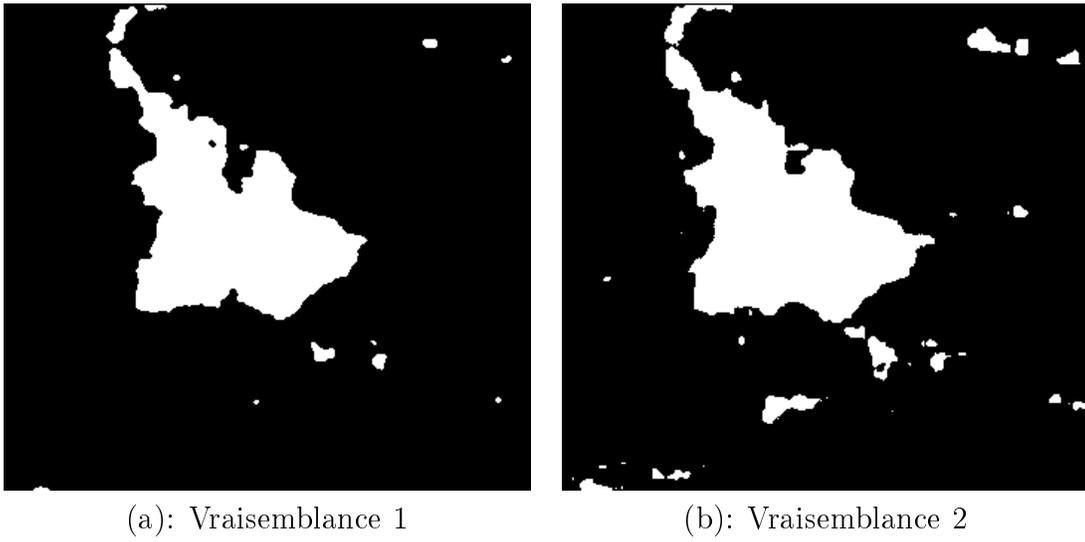
L'approche globale proposée et schématisée figure 3.22 sous la forme de boîtes noires s'adapte à d'autres types de problèmes de segmentation/classification à condition d'adapter les étapes d'extraction et de fusion des paramètres afin de ne retenir que les plus pertinents d'entre eux pour le problème de traitement d'image à traiter.

Nom Image Taille	Capteur Résolution	Taille Fenêtre	CPU texture	CPU fuzzy Cmeans	CPU ICM
La Roche sur Yon 512×512	SPOT3 10m	11×11	3min54s	3min58s	43s
Cayenne 1078×667	SPOT3 10m	13×13	11min22s	14min12s	5min14s
Epeville 1003×900	SPOT5 5m	23×23	17min4s	8min35s	1min10s
Toreilles 377×350	SPOT5 5m	17×17	2min4s	2min4s	10s
Canet en Roussillon 423×407	SPOT5 5m	17×17	2min46s	2min31s	18s
Toulouse 1006×1178	SPOT5 2m50	27×27	24min37s	15min20s	5min45s
Belém 592×1024	ERS1 (8bits) 12m50	15×15	10min	12min47s	2min40s
Macapá 1197×1024	ERS1 (8bits) 12m50	15×15	15min	16min10s	4min
Macapá 1370×1270	ERS1 (16 bits) 12m50	15×15	3h15min	22min30s	8min
Aix en Provence 1024×1024	ERS1 (16 bits) 12m50	15×15	2h8min	13min35s	5min20s

**Tableau 3.2:** Valeurs des paramètres et temps de calcul sur station SUN ultra1 (124MB, 167 MHz).



**Figure 3.9:** *Canet en Roussillon : simulation SPOT5 à 5m.*



**Figure 3.10:** *Toreilles : simulation SPOT5 à 5m.*

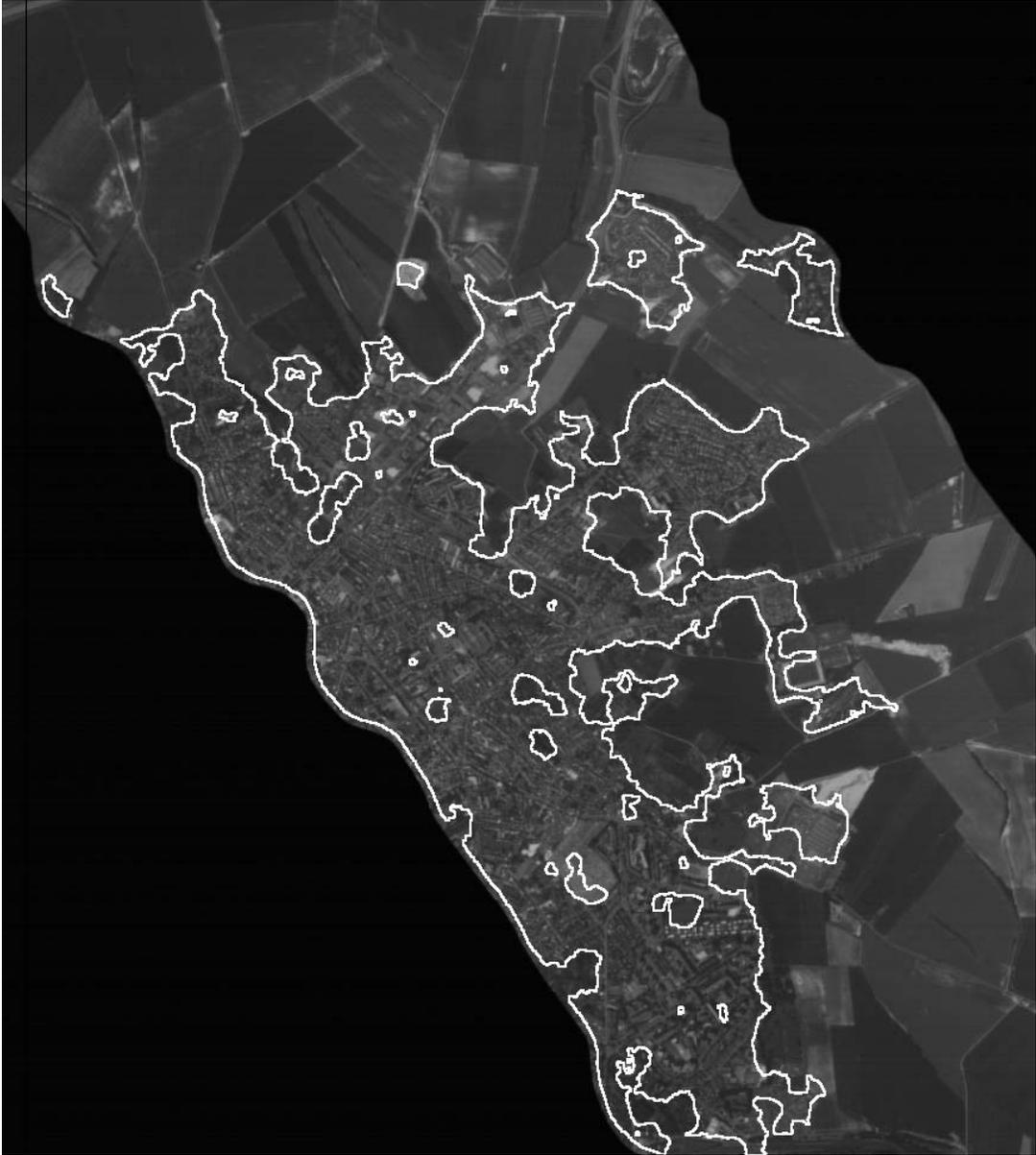


Figure 3.11: *Epperville : simulation SPOT5 à 5m.*



**Figure 3.12:** *Canet en Roussillon : simulation SPOT5 à 5m.*



**Figure 3.13:** *Toreilles : simulation SPOT5 à 5m.*

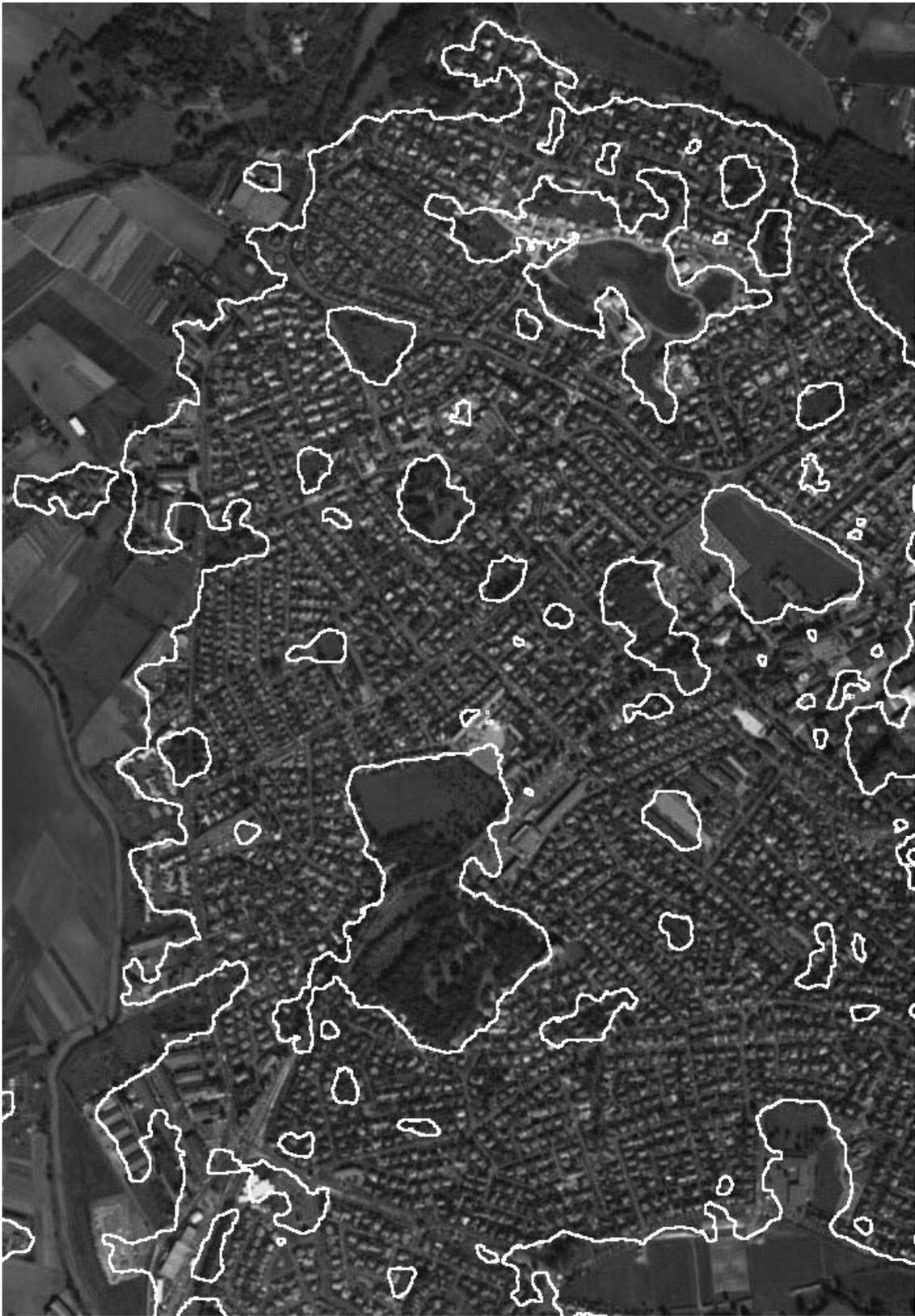


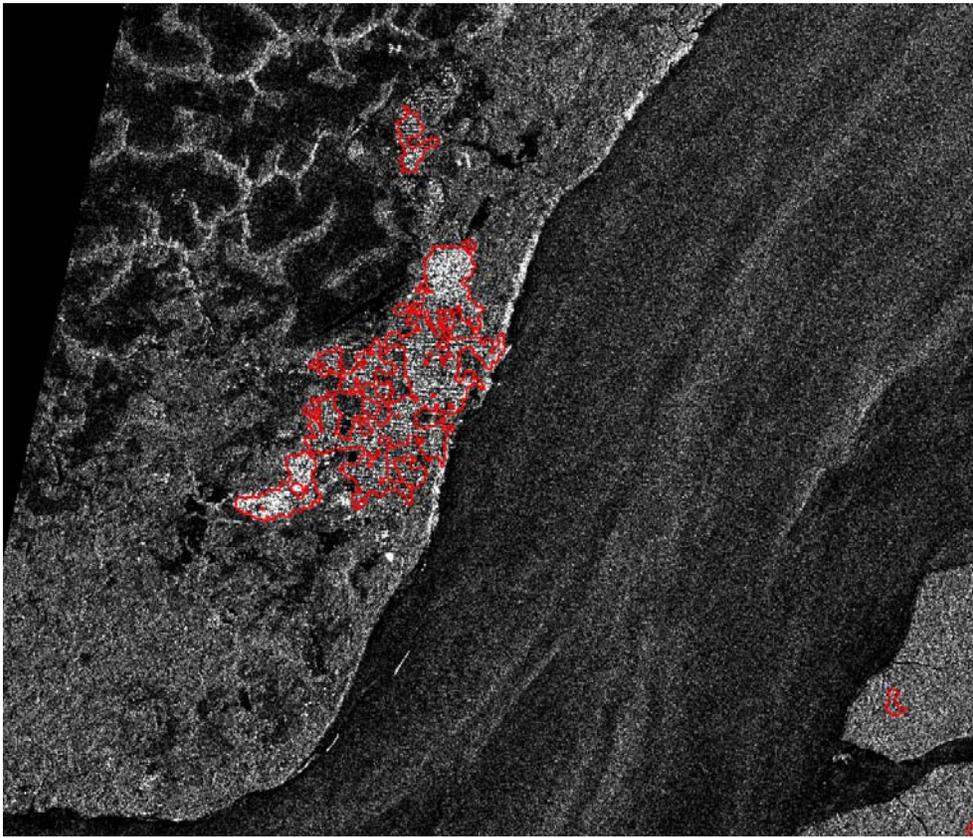
Figure 3.14: *Toulouse : simulation SPOT5 à 2m50.*



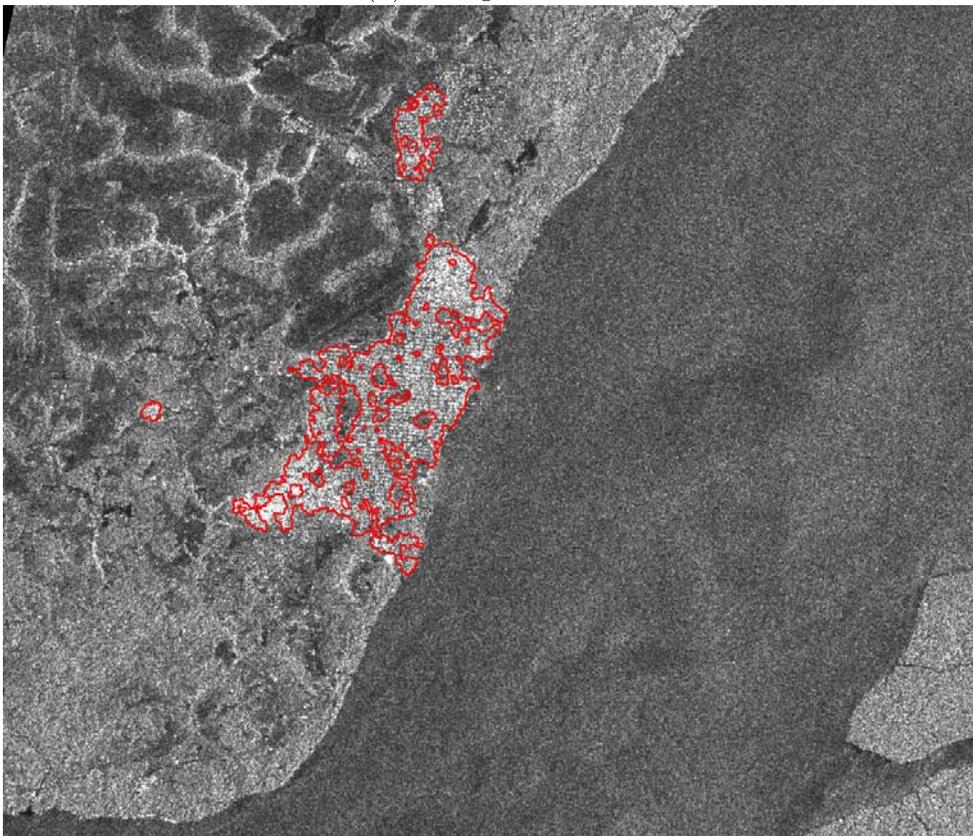
Figure 3.15: *La Roche sur Yon* : image SPOT3 à 10m.



**Figure 3.16:** Cayenne : image SPOT3 à 10 m.

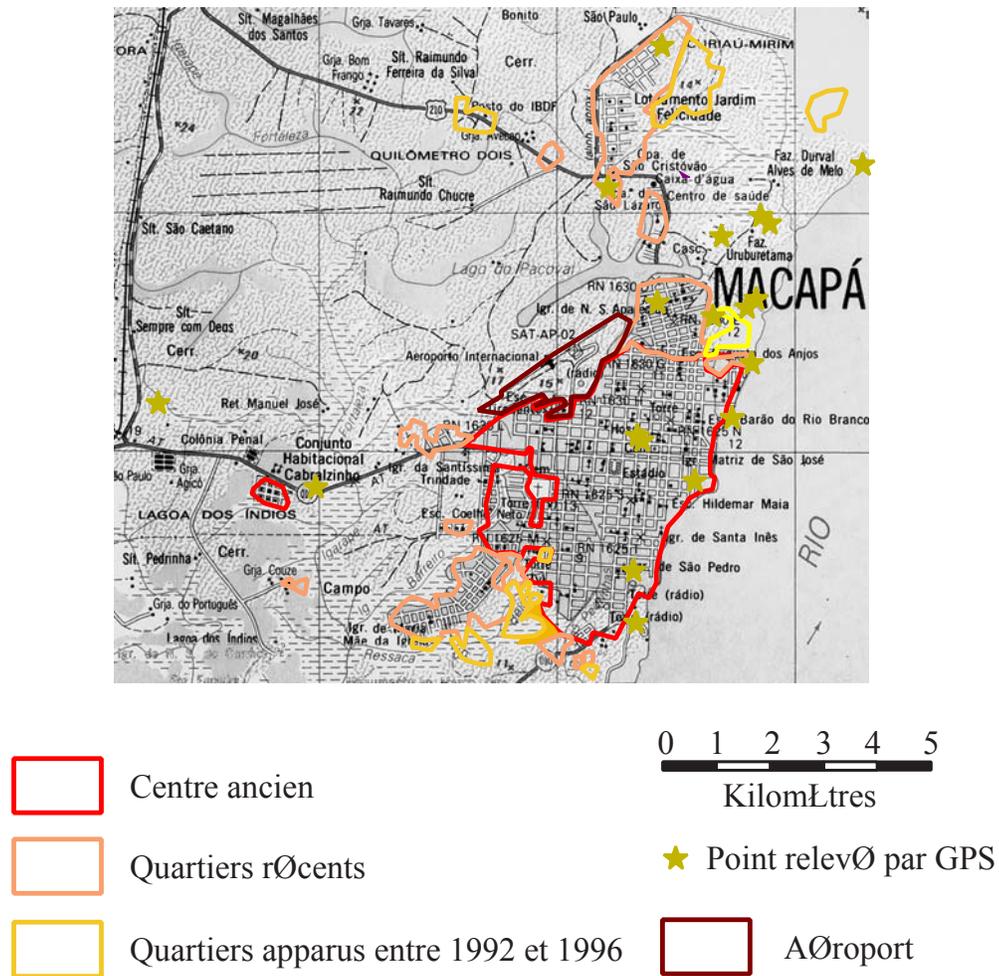


(a): Image de 1992



(b): Image de 1996

**Figure 3.17:** *Macapá (Brésil) : images ERS1 à 12m50 (8 bits).*



**Figure 3.18:** Projection des limites des zones urbaines obtenues par photo-interprétation d'images radar sur la carte topographique (1:100 000, 1995) fournie par l'Institut Brésilien de Géographie et de Statistique.

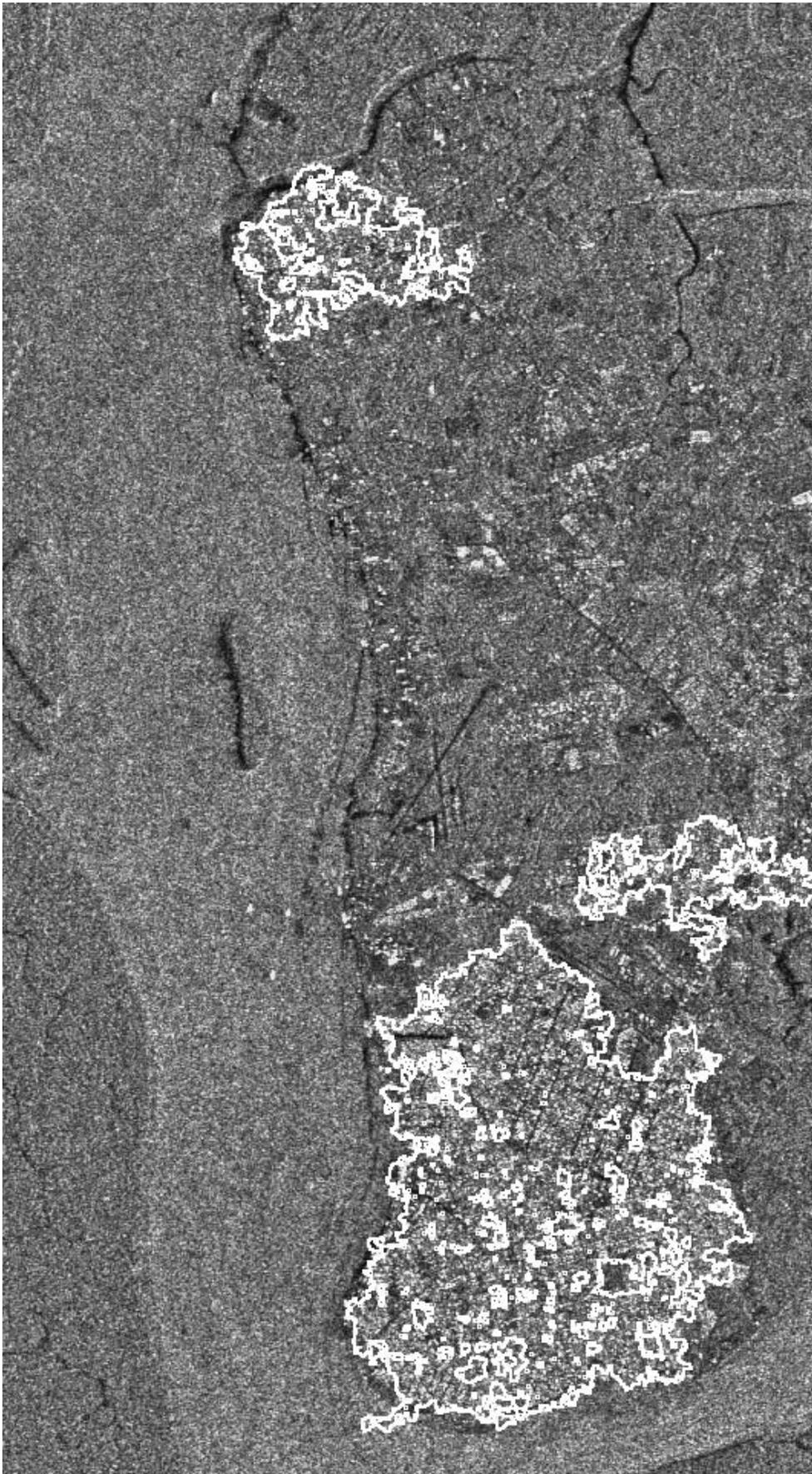
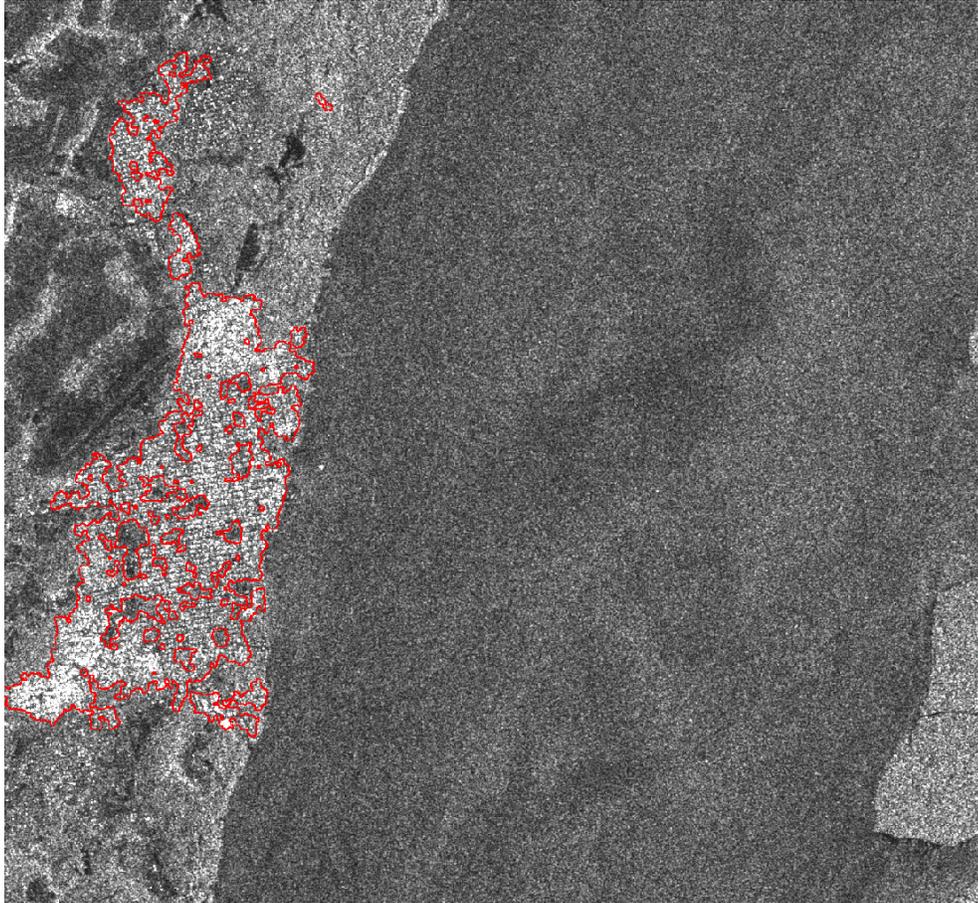
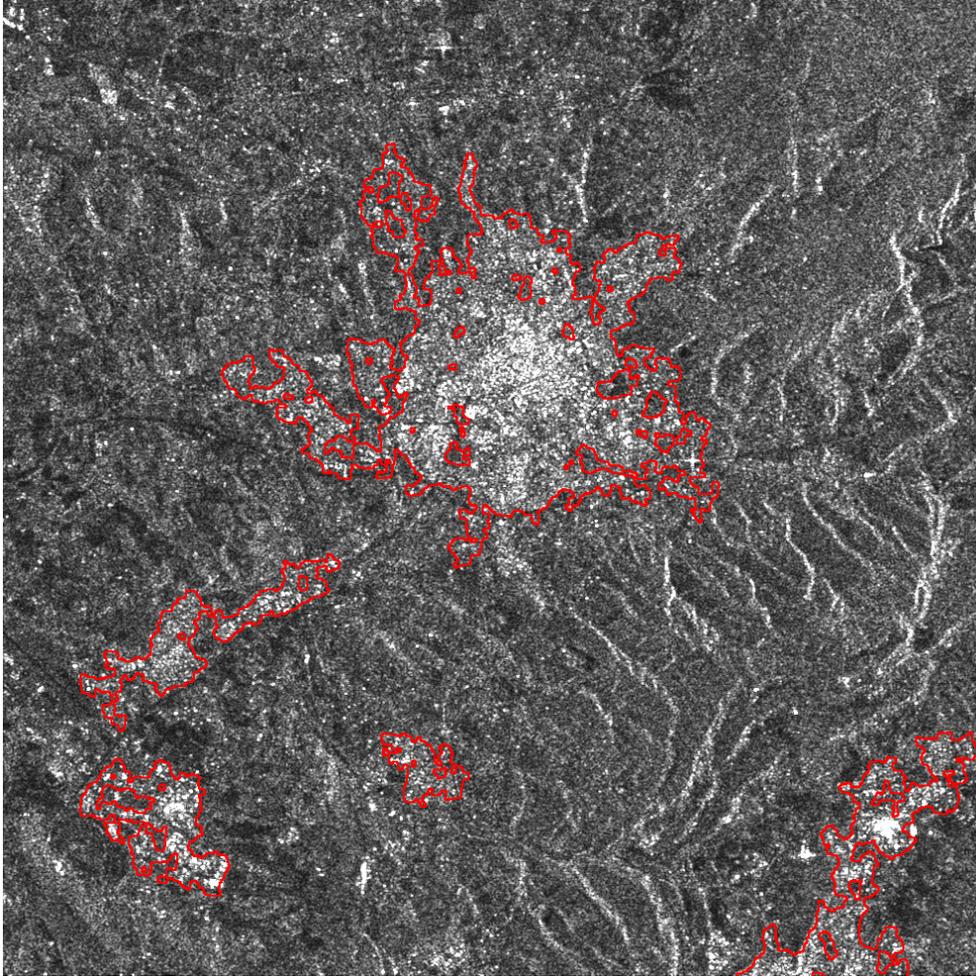


Figure 3.19: *Belém (Brésil) : image ERS1 à 12m50 (8 bits).*



**Figure 3.20:** *Macapá (Brésil) : image ERS1 à 12m50 (16 bits).*



**Figure 3.21:** Aix en Provence : image ERS1 à 12m50 (16 bits).

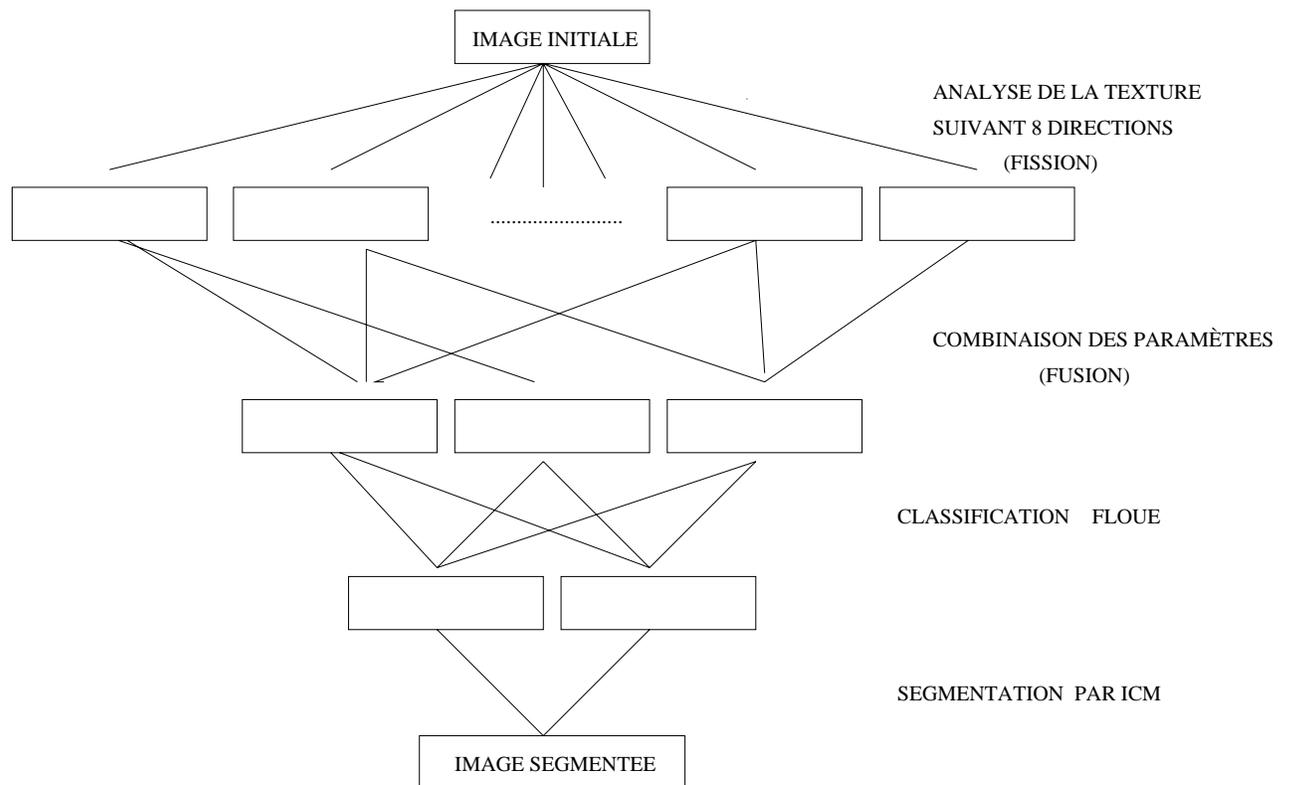


Figure 3.22: Schéma global.

## DANS CE CHAPITRE:

4.1	Introduction . . . . .	112
4.2	Analyse intra-urbaine: quelques applications . . . . .	112
4.3	Généralités et définitions . . . . .	114
4.3.1	Morphologie mathématique : quelques généralités . . . . .	114
4.3.2	Les transformations de base : définitions et propriétés . . . . .	115
4.3.2.1	Définitions . . . . .	115
4.3.2.2	Propriétés algébriques . . . . .	118
4.3.2.3	Propriétés géométriques . . . . .	119
4.3.3	Ouverture par Reconstruction . . . . .	119
4.3.3.1	Introduction . . . . .	119
4.3.3.2	Transformations géodésiques . . . . .	120
4.3.3.3	Ouverture par reconstruction . . . . .	121
4.3.3.4	Ouverture par reconstruction versus ouverture morphologique : exemples dans les cas binaire et en niveaux de gris . . . . .	123
4.4	Granulométrie et distribution de taille . . . . .	124
4.4.1	Introduction . . . . .	124
4.4.2	Granulométrie . . . . .	127
4.4.3	Distribution de taille . . . . .	128
4.4.4	Application à l'analyse intra-urbaine . . . . .	129
4.4.4.1	Granulométrie . . . . .	129
4.4.4.2	Quantification . . . . .	131
4.5	Modèle markovien de segmentation . . . . .	131
4.5.1	Introduction . . . . .	131
4.5.2	Le modèle . . . . .	132
4.5.3	Algorithme du recuit simulé . . . . .	133
4.5.4	Segmentation : résultats expérimentaux . . . . .	134
4.6	Conclusion . . . . .	137

# 4.

## Analyse Intra-Urbaine

L' extraction d'un masque urbain est bien souvent une étude préalable à l'analyse intra-urbaine. Dans ce chapitre, nous nous proposons d'étudier dans quelle mesure l'analyse granulométrique, issue de la théorie de la morphologie mathématique, permet de caractériser le tissu urbain. Cette approche consiste à appliquer une série de transformations morphologiques à l'aide d'éléments structurants de taille croissante. Nous présentons donc dans un premier temps les transformations de base de la morphologie mathématique avant de nous focaliser sur l'ouverture par reconstruction qui est la transformation de base de la granulométrie utilisée lors de ce travail.

## 4.1 Introduction

Avec l'augmentation de la résolution des images satellitaires, les notions de taille, de forme d'objets deviennent les paramètres discriminants pour l'analyse intra-urbaine. La morphologie mathématique et plus particulièrement la granulométrie, offre un cadre de travail intéressant pour analyser ces caractéristiques.

L'étude présentée ici a pour but essentiel d'étudier les potentialités de cette approche pour analyser la texture intra-urbaine afin de la segmenter. En effet, la granulométrie a été largement appliquée dans les domaines de la géologie [110], du biomédical [112] mais très peu dans le domaine de la télédétection.

Après avoir donné quelques exemples d'utilisation d'une classification urbaine, nous introduisons la morphologie mathématique. Les transformations de base sont définies dans le cas binaire et dans le cas des images en niveaux de gris. La transformation, moins classique, d'ouverture par reconstruction, utilisée pour notre application, est présentée. Elle est fondée sur des transformations géodésiques qui sont elles même définies.

Après toute transformation, il est nécessaire de quantifier les objets que l'on souhaite caractériser afin de les classer. Il s'agit donc d'extraire des paramètres pertinents représentant une information simplifiée. Dans cette optique, nous présentons l'approche par granulométrie. Les paramètres extraits par cette méthode sont ensuite introduits dans un modèle markovien de classification.

L'algorithme global est appliqué à des simulations SPOT5.

## 4.2 Analyse intra-urbaine: quelques applications

Dans le cadre des projets EUROSTAT et ATLAS, présentés dans 1.1, des études ont été réalisées sur certaines villes (e.g. Toulouse, Berlin, Athènes) afin d'extraire plusieurs classes dans le tissu urbain. Soulignons que, dans ce cas, des données exogènes ont été nécessaires pour extraire certaines classes de la nomenclature CLUSTER (définie dans le cadre du projet EUROSTAT) qui mélange occupation du sol et usage du sol. Lorsque l'objectif est d'extraire quelques classes dans le tissu urbain à partir de la seule donnée satellitaire, l'approche par analyse de la texture est très souvent utilisée. Pour un état de l'art dans ce domaine le lecteur se référera au paragraphe 1.2.3. Les deux étapes d'extraction du masque urbain et d'analyse intra-urbaine sont la plupart du temps réalisées dans une seule et même étape. Cependant Gouinaud [47] montre qu'une approche en deux étapes (détection du masque puis analyse intra-urbaine) est bien adaptée au milieu urbain et permet d'améliorer les résultats.

Nous présentons dans la suite de ce chapitre deux domaines pour lesquels une analyse intra-urbaine est nécessaire. Elle est nécessaire pour les études préalables à l'implantation d'un nouveau réseau de téléphonie cellulaire [101]. En effet, la propagation des ondes dépend de la nature du sol. A l'heure actuelle, ces plans d'occupation du sol sont réalisés par photo-interprétation d'images satellitales et photo-interprétation assistée par ordinateur (PIAO). La société ISTAR fournit notamment à ces clients (fournisseurs de réseaux de téléphonie mobile) des produits appelés *clutter map* réalisés par PIAO.

Le plan d'occupation du sol en milieu urbain sert également de base pour des sondages démographiques. Dans le cahier du GDTA [4], les auteurs présentent un exemple d'étude sur la ville de Quito. En effet, dans les pays émergents, les méthodes classiques d'observation des populations citadines (recensements exhaustifs, enquêtes par sondage) sont difficilement applicables par manque d'informations cartographiques et de bases de sondage. De plus, de par leur coût et leur durée d'exploitation, elles ne sont pas adaptées à l'étude des populations à croissance et évolution rapides. L'image satellitale offre une alternative intéressante. La classification de l'image est réalisée suivant la densité du bâti. Cette information est ensuite utilisée afin de sélectionner des échantillons pour une enquête démographique sur le terrain.

Dans le but de mieux gérer l'espace urbain, le Centre National de Télédétection (CNT), le District de Tunis (DT), et la Direction Générale de l'Aménagement du Territoire et de l'Urbanisme (DGATU) ont conduit un projet pilote [4] sur la région du grand Tunis. Une des étapes importantes de cette étude a été la réalisation d'une carte d'occupation du sol de la ville de Tunis en 1990 à l'échelle 1/100 000. Cette carte a été réalisée à partir de données satellitales SPOT (XS et P). Les méthodes employées sont à la fois des méthodes automatiques d'analyse de texture et des méthodes de PIAO.

En ce qui concerne les applications concrètes de l'utilisation du traitement d'images satellitales, la gestion du paysage urbain semble occuper une place prédominante.

## 4.3 Généralités et définitions

### 4.3.1 Morphologie mathématique : quelques généralités

Le terme "morphologie", qui vient des termes grecs *morphê* qui signifie forme et *logos* qui signifie science, traduit l'idée de base de la théorie de l'analyse de la forme des objets. Cette théorie a été fondée à la fin des années 60 par Matheron [52, 53, 89] et Serra [110, 111, 112] afin d'analyser les structures spatiales de l'image que l'on étudie en la comparant localement avec un élément de base, appelé élément structurant, dont on maîtrise la forme et la taille. La comparaison se fait à partir de relations ensemblistes (inclusion, intersection,  $\dots$ ). C'est pourquoi en traitement d'image, les images, qu'elles soient binaires ou en niveaux de gris, sont décrites comme des ensembles comme nous le verrons par la suite.

Dans [110], Serra explique comment la théorie a été développée dans le but de répondre à des problèmes aussi divers et éloignés que : l'étude de la dynamique du mouvement des nuages, la reconnaissance d'écriture, la détection de cellules cancéreuses à partir d'un frottis  $\dots$ . L'idée de départ a été de rechercher les dénominateurs communs entre tous ces problèmes. En fait, dans les sciences exactes, le scientifique opère toujours en deux phases distinctes : la première étape consiste à transformer l'objet que l'on souhaite étudier afin de le simplifier, de l'isoler, la seconde étape consiste à le quantifier, le mesurer à l'aide de paramètres pertinents qui permettent de caractériser cet objet par des propriétés physiques (e.g. vitesse de propagation du son dans une roche, résistance de la roche aux chocs [110]).

Il faut souligner que la morphologie mathématique [106] est une des rares approches développées dès le départ pour le traitement d'image et appliquée par la suite au traitement du signal. Le développement de la théorie a toujours été motivé par des applications pratiques. La première application a été l'étude du minerai de fer lorrain [109] puis la caractérisation des milieux poreux [88]. Elle a été, par la suite, appliquée à de nombreux problèmes de traitement d'image tels que le filtrage, la segmentation, la description de formes, le codage [106]. La morphologie mathématique permet également l'analyse de texture [67]. Dans [123], Voiron montre la pertinence de la morphologie mathématique pour l'analyse des organisations et des dynamiques spatiales en géographie.

### 4.3.2 Les transformations de base : définitions et propriétés

Les transformations de base de la morphologie mathématique sont la dilatation et l'érosion. L'ouverture et la fermeture sont définies à partir d'elles. Ces opérateurs sont non linéaires et non inversibles. D'abord appliquées au cas d'images binaires, ces transformations ont été ensuite étendues pour traiter des images en niveaux de gris. Les opérateurs sont définis pour des images définies sur  $\mathbb{R}^2$ . Le problème de la digitalisation a été étudié dans [106, 111].

#### 4.3.2.1 Définitions

Avant toute chose, nous définissons l'addition (notée  $\oplus$ ) et la soustraction (notée  $\ominus$ ) de Minkowski d'un ensemble  $A$  par un ensemble  $B$  utilisée à l'origine en géométrie intégrale [91] :

$$\begin{aligned} A \oplus B &= \{a + b \mid a \in A \text{ et } b \in B\} \\ A \ominus B &= \{u \mid \forall b \in B, b + u \in A\} \end{aligned}$$

Dans le cas binaire, l'image est décrite par l'ensemble  $A \subset \mathbb{R}^2$ . Dans ce qui suit, nous notons  $B$  l'élément structurant,  $\check{B} = \{b \mid -b \in B\}$  le symétrique de  $B$  par rapport à l'origine et  $B_s = \{s + t, t \in B\}$  le translaté de  $B$  en  $s$ .

Soulignons que dans le cas d'un élément structurant symétrique, e.g. un disque,  $\check{B}$  et  $B$  sont identiques. Dans le tableau 4.1 les transformations de base sont définies dans le cas d'images binaires.

Nous avons appliqué chacune de ces transformations à l'image binaire présentée figure 4.1 (a). Les résultats sont présentés figures 4.1 (b), (c), (d) et (e).

Pour étendre ces opérateurs aux cas des images en niveaux de gris, Serra [110, 111] définit la section supérieure d'une fonction  $f$  au niveau  $g$  par :

$$X_g(f) = \{s \mid f(s) \geq g\}$$

Une image en niveaux de gris est aussi caractérisée par son ombre, appelée également sous-graphe :

$$SG(f) = \{(s, g) \in \mathbb{R}^2 \times \mathbb{R} \mid f(s) \geq g\} = \{(s, g) \in \mathbb{R}^2 \times \mathbb{R} \mid s \in X_g(f)\}$$

Serra définit alors les fonctions  $f$  semi-continues supérieure<sup>1</sup> définies sur  $\mathbb{R}^2$  et à valeurs dans  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ . Par convention,  $f(s) = -\infty$  si  $s$  n'appartient

<sup>1</sup> $f$  est semi-continue supérieure ssi  $\forall s, \forall g > f(s), \exists$  un voisinage  $V(s)$  tq  $f(t) < g, \forall t \in V(s)$ .



(a): objet initial



(b): objet dilaté



(c): objet érodé



(d): objet ouvert



(e): objet fermé

**Figure 4.1:** Transformations de base (dilatation, érosion, ouverture, fermeture) par un disque de diamètre 17.

nom	notation	définition
dilatation	$A \oplus \check{B}$	$\{s   B_s \cap A \neq \emptyset\}$
érosion	$A \ominus \check{B}$	$\{s   B_s \subseteq A\}$
ouverture	$A_B$	$(A \ominus \check{B}) \oplus B$
fermeture	$A^B$	$(A \oplus \check{B}) \ominus B$

**Tableau 4.1:** Transformations de base dans le cas binaire.

pas au support de  $f$ .

Une image en niveaux de gris est alors une fonction positive, semi-continue supérieure<sup>2</sup>, bornée :  $0 \leq f(s) \leq N$ ,  $\forall s \in \mathbb{R}^2$ . A noter que dans ce cas, le support de  $f$  est l'ensemble des  $s$  tel que  $f(s) \neq 0$ .

Notons que  $SG(f)$  est un fermé ssi  $f$  est semi-continue supérieure. Les images en niveaux de gris, bornées, vérifient cette propriété.

Pour définir les opérateurs morphologiques dans le cas des images en niveaux de gris (présentés dans le tableau 4.2), il faut appliquer aux ombres les opérateurs définis précédemment sur les images binaires. Nous notons  $\vee$  le sup,  $\wedge$  l'inf, et  $\check{h}(t) = h(-t)$ .

L'ouverture, notée  $f_h$  (resp. la fermeture, notée  $f^h$ ) est définie comme une érosion suivie d'une dilatation (resp. une dilatation suivie d'une érosion).

Si l'on travaille avec un élément structurant plan<sup>3</sup>  $h_K(t)$ , défini sur un compact  $K \subset \mathbb{R}^2$ , les expressions précédentes (Tableau 4.2) se simplifient de la manière

<sup>2</sup>L'hypothèse de semi-continuité supérieure de  $f$  est nécessaire et suffisante pour que  $SG(f)$  soit un fermé.

<sup>3</sup> $h_K(t)$  vaut 0 sur  $K$  et  $-\infty$  ailleurs.

nom	notation	définition
dilatation	$(f \oplus \check{h})(t)$	$\bigvee_{s \in \mathbb{R}^2} [f(s) + h(s - t)]$
érosion	$(f \ominus \check{h})(t)$	$\bigwedge_{s \in \mathbb{R}^2} [f(s) - h(s - t)]$

**Tableau 4.2:** Transformations de base dans le cas des images en niveaux de gris.

suivante :

$$\begin{aligned} f \oplus \check{h}_K(t) &= \bigvee_{s \in K_t} f(s) \text{ et } (f \oplus \check{h}_K)_g = X_g(f) \oplus \check{K}. \\ f \ominus \check{h}_K(t) &= \bigwedge_{s \in K_t} f(s) \text{ et } (f \ominus \check{h}_K)_g = X_g(f) \ominus \check{K}. \end{aligned}$$

Ces relations montrent que, dans le cas d'un élément structurant plan, on peut effectuer les dilatations et les érosions sur les sections supérieures indépendamment.

#### 4.3.2.2 Propriétés algébriques

Que ce soit dans le cas binaire ou dans le cas des images en niveaux de gris les transformations d'ouverture et de fermeture vérifient les propriétés algébriques suivantes :

1. croissance,  $\forall X, Y \subset \mathbb{R}^n, X \subset Y \Rightarrow X_B \subset Y_B$  et  $X^B \subset Y^B$ .
2. anti-extensivité pour l'ouverture, extensivité pour la fermeture,  $\forall X \subset \mathbb{R}^n, X_B \subset X$  et  $X \subset X^B$ .
3. idempotence,  $\forall X \subset \mathbb{R}^n, (X_B)_B = X_B$  et  $(X^B)^B = X^B$ .

$n = 2$  dans le cas d'images binaires et  $n = 3$  dans le cas d'images en niveaux de gris.

Notons que Matheron [89] définit l'ouverture algébrique et la fermeture algébrique comme étant des opérateurs vérifiant les propriétés définies ci-dessus.

### 4.3.2.3 Propriétés géométriques

Les ouvertures et fermetures morphologiques sont des ouvertures algébriques qui vérifient les 4 principes de la morphologie mathématique nécessaires à l'étape de quantification dans  $\mathbb{R}^n$  [110] :

1. invariance par translation :  $\psi$  est invariante par translation si translater  $X$  par  $h$  et ensuite appliquer  $\psi$  est équivalent à appliquer  $\psi$  et ensuite translater le résultat par  $h$ , i.e.  $\psi(X_h) = [\psi(X)]_h$ .
2. compatibilité par changement d'échelle : la transformation  $\psi$  vérifie  $\psi_\lambda(X) = \lambda\psi\left(\frac{X}{\lambda}\right)$ .
3. connaissance locale :  $\psi$  vérifie ce principe si pour tout ensemble borné  $E$  dans lequel nous voulons connaître  $\psi(X)$ , on peut trouver un ensemble borné  $F$  dans lequel la connaissance de  $X$  suffit pour appliquer localement (i.e. dans  $E$  la transformation).
4. semi-continuité : les contours de la transformée sont égaux à la transformée des contours.

Par la suite, d'autres transformations plus complexes ont été proposées telles que les filtres séquentiels alternés, la transformation du chapeau haut de forme, les opérateurs connexes.

## 4.3.3 Ouverture par Reconstruction

### 4.3.3.1 Introduction

L'ouverture par reconstruction appartient à un ensemble plus général de transformations appelées opérateurs connexes [105]. C'est d'ailleurs le premier opérateur connexe connu. Le terme "reconstruction" traduit une phase de reconstruction des composantes connexes de l'image. Notons que ce terme s'applique plus généralement à toutes les opérations classiques précédemment définies i.e. la fermeture, les filtres alternés séquentiels. Selon Vincent [122], elle permet de traiter des problèmes variés du traitement d'image tels que le filtrage, la segmentation, etc.

Les opérateurs connexes ont la propriété fondamentale de simplifier l'image tout en préservant les contours contrairement aux opérateurs classiques précédemment définis qui dégradent les contours des objets.

Dans le cas binaire la définition d'un opérateur connexe est la suivante :

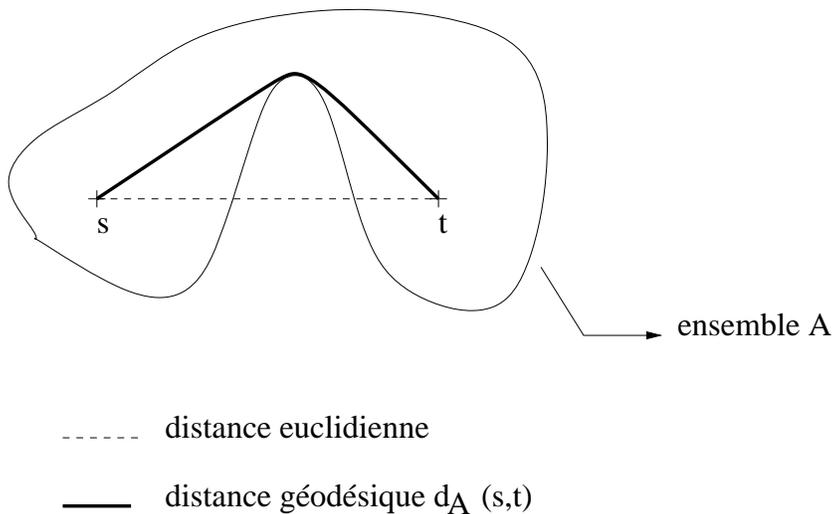
**Définition 4.3.1** *Un opérateur  $\psi$  qui s'applique sur des images binaires est connexe si pour toute image binaire  $X$ ,  $X \setminus \psi(X)$  est exclusivement composé de composantes connexes de  $X$  ou de son complément  $X^c$ .*

L'extension de la définition au cas des images en niveaux de gris peut se faire comme nous le verrons dans la suite, i.e. dans le cas particulier de l'ouverture par reconstruction, à partir de la décomposition de l'image par seuillages successifs auxquels on applique l'opérateur binaire. Cette phase est suivie d'une phase d'empilement [57] afin de reconstruire l'image en niveaux de gris transformée.

L'ouverture par reconstruction [121] se définit à partir des transformations géodésiques que nous présentons dans la suite.

#### 4.3.3.2 Transformations géodésiques

Les transformations géodésiques sont fondées sur la définition d'un écart, appelé distance géodésique par abus de langage. Cette métrique a été définie pour répondre à certains besoins comme l'étude des propriétés de connexité.



**Figure 4.2:** *Distance géodésique versus distance euclidienne.*

**Définition 4.3.2** *Étant donné un ensemble  $A \subset \mathbb{R}^2$ , la distance géodésique<sup>4</sup> entre deux pixels  $s$  et  $t$  dans  $A$ , notée  $d_A(s, t)$ , est la borne inférieure des longueurs des chemins<sup>5</sup> joignant  $s$  et  $t$ , totalement inclus dans  $A$ .*

A partir de cette distance, on définit la dilatation géodésique [72] de taille  $r \geq 0$ .

**Définition 4.3.3** *La dilatation géodésique de taille  $r$  d'un ensemble  $A$  dans un ensemble  $D$ , tel que  $D \subset \mathbb{R}^2$  et  $A \subseteq D$ , est l'ensemble des pixels de  $D$  dont la distance géodésique à  $A$  est inférieure ou égale à  $r$  :*

$$\delta_D^{(r)}(A) = \{s \in D \mid d_D(s, A) \leq r\}.$$

La dilatation géodésique de taille  $r$  est obtenue en itérant  $r$  fois la dilatation géodésique élémentaire :

$$\delta_D^{(r)}(A) = \underbrace{\delta_D^{(1)}(A) \circ \delta_D^{(1)}(A) \cdots \circ \delta_D^{(1)}(A)}_{r \text{ fois}}.$$

La dilatation géodésique élémentaire, qui est une dilatation conditionnelle, est définie comme suit :

$$\delta_D^{(1)}(A) = (A \oplus B_1) \cap D$$

$B_1$  est l'élément structurant élémentaire associé au réseau de pixels.  $B_1$  est donc défini sur la figure 4.3 suivant la connexité choisie sur une trame carrée.

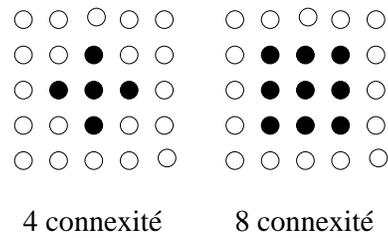
### 4.3.3.3 Ouverture par reconstruction

Nous pouvons maintenant définir la reconstruction à partir de la dilatation géodésique.

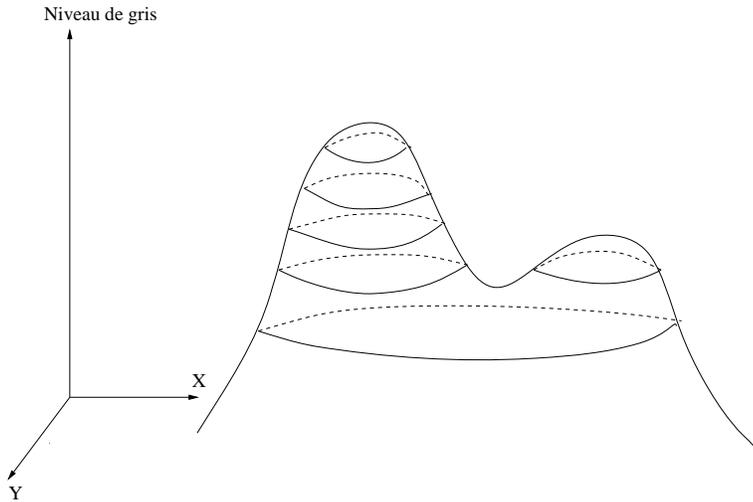
**Définition 4.3.4** *La reconstruction de  $D$  à partir de  $A \subseteq D$  est définie comme suit :  $\rho_D(A) = \lim_{r \rightarrow \infty} \delta_D^{(r)}(A)$ .*

<sup>4</sup>la distance géodésique est en fait un écart car elle peut prendre des valeurs infinies.

<sup>5</sup>Un chemin d'extrémités  $s$  et  $t$  dans  $A$  est une application continue  $c : [a, b] \rightarrow A$  telle que  $c(a) = s$  et  $c(b) = t$ .



**Figure 4.3:** *Élément structurant unité défini en 4-connexité et 8-connexité.*



**Figure 4.4:** *Sections supérieures.*

$A$  est appelée image de marqueurs et  $D$  le masque. Cette transformation consiste donc à conserver toutes les composantes connexes  $D$  qui contiennent au moins un pixel de  $A$ .

Dans [106, 122], les auteurs étendent cette définition au cas des images en niveaux de gris en utilisant la décomposition de l'image par seuillages [57] successifs (Figure 4.4). Supposons que l'image  $f$  soit à valeurs dans  $[0, N - 1]$ . A condition d'utiliser des éléments structurants plan, on peut travailler sur les différentes sections supérieures de l'image. La reconstruction de  $f$  dans  $h$  est alors définie comme suit :

$$\rho_h(f)(s) = \max\{g \in [0, N - 1] \mid s \in \rho_{X_g(h)}(X_g(f))\}.$$

La dilatation géodésique de taille  $r \geq 0$  de  $f$  dans  $h$  est alors définie de la manière

suivante :

$$\delta_h^{(r)}(f) = \underbrace{[(f \oplus B_1) \wedge h] \circ [(f \oplus B_1) \wedge h] \cdots \circ [(f \oplus B_1) \wedge h]}_{r \text{ fois}}.$$

Une autre définition de la reconstruction de  $f$  dans  $h$ , découle de la définition précédente :

$$\rho_h(f) = \lim_{r \rightarrow \infty} \delta_h^{(r)}(f)$$

En pratique, nous effectuons une série de dilatations géodésiques jusqu'à stabilisation. Dans la suite on parlera d'ouverture par reconstruction ( $\rho_f(h) = \rho_f(f_B)$ ) car l'image de marqueurs  $h$  est obtenue par ouverture de l'image originale  $f$  par un élément structurant  $B$ , i.e.  $h = f_B$ . Cette image de marqueurs est ensuite dilatée géodésiquement. Il est également possible de définir des fermetures géodésiques. Enfin soulignons que l'ouverture par reconstruction [106] est une ouverture algébrique, i.e. une opération qui vérifie les propriétés de croissance, d'anti-extensivité, et d'idempotence.

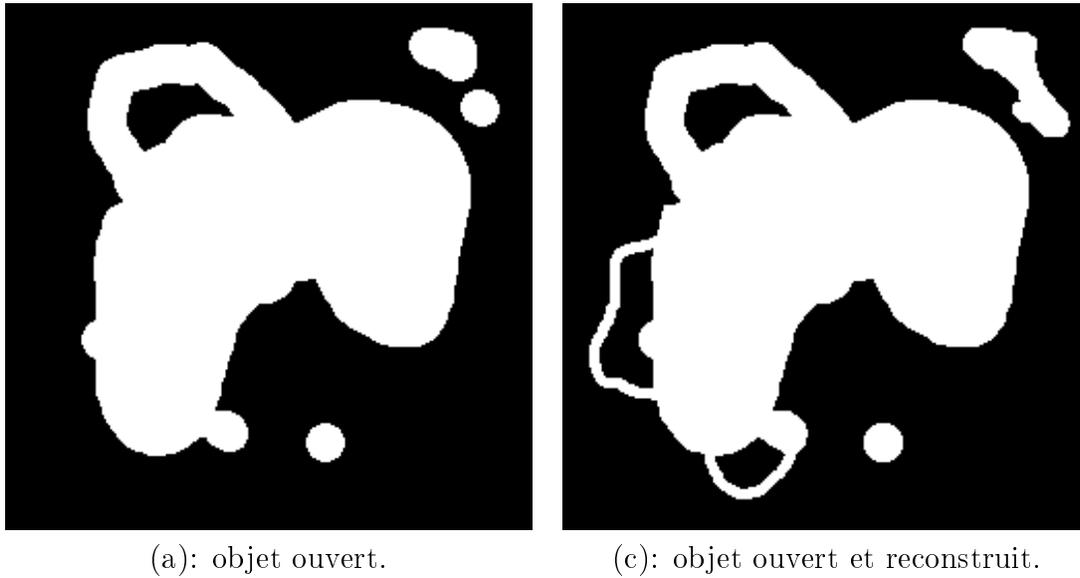
L'ouverture par reconstruction a été utilisée pour des applications aussi diverses que la détection de mines [7], l'extraction de micro-anévrismes [122].

#### 4.3.3.4 Ouverture par reconstruction versus ouverture morphologique : exemples dans les cas binaire et en niveaux de gris

Une ouverture classique non suivie d'une reconstruction a pour conséquence une modification des contours des objets binaires. Si l'on applique l'ouverture par reconstruction, les objets qui ne sont pas filtrés lors de l'ouverture sont conservés intacts (Figure 4.5 (b) versus (c)).

De même dans le cas des images en niveaux de gris (Figure 4.6), les contours après transformation correspondent à certains contours de l'image originale. Sur l'image 4.6 (c) la transformation a créé des zones plates. Cette transformation n'est pas sans rappeler l'approche proposée par Gouinaud dans [47] qui utilise le filtre de Nagao afin de créer dans son image des zones plates, i.e. "de supprimer les détails de l'image de taille inférieure aux structures intéressantes de l'habitat".

La reconstruction est intéressante dans la mesure où les objets, ou structures que l'on souhaite caractériser sont bien séparés. Dans le cas contraire la reconstruction peut avoir des effets pervers (Figure 4.7). En effet, la reconstruction permet



**Figure 4.5:** Ouverture et ouverture par reconstruction de l'image présentée 4.1 (a) par un disque de diamètre 17 .

de retrouver des contours présents dans l'image originale mais si les structures ne sont pas bien séparées, quelle signification ont ces contours ? Dans [105], l'auteur souligne le problème de "fuite" (*leakage* en anglais) des opérateurs connexes, i.e. certains petits objets qui devraient être supprimés par la transformation ne le sont pas totalement car ils fusionnent entre eux et/ou sont connectés avec les régions qui les entourent.

Nous nous plaçons dans l'optique où les images sur lesquelles nous travaillons vont avoir une résolution de plus en plus grande (cf. § 1.1) et donc sur lesquelles les structures seront plus nettement séparées.

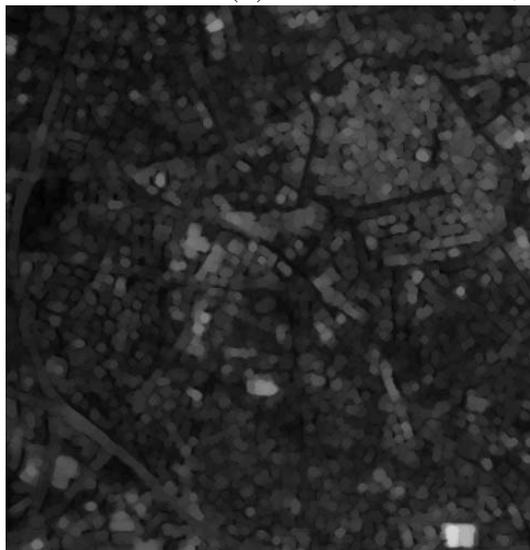
## 4.4 Granulométrie et distribution de taille

### 4.4.1 Introduction

La granulométrie permet d'analyser des objets, de filtrer des images suivant des critères de forme. Elle a été utilisée largement dans des domaines aussi variés que la géologie [111], l'imagerie médicale [49], l'analyse de texture [31, 32, 115], la



(a): Aix en Provence, simulation SPOT5 à 5m.



(b): Ouverture.



(c): Ouverture  
par Reconstruction.

**Figure 4.6:** *Ouverture (CPU=0s57) et ouverture par reconstruction (CPU=3s89) par un disque de diamètre 7 de l'image d'Aix en Provence.*



(a): La Roche sur Yon, image SPOT3 à 10m.



(b): Ouverture  
par Reconstruction.

**Figure 4.7:** *Ouverture par reconstruction (CPU=8s47) par un disque de diamètre 7 de l'image de La Roche sur Yon.*

détection d'objets [87]. Toutefois, elle est plus rarement appliquée dans le domaine de la télédétection. Avec l'augmentation de la résolution des images satellitales, les critères de taille et de forme des objets vont devenir prépondérants. Il est donc intéressant d'utiliser une telle approche permettant de distinguer des zones dans la ville en tenant compte de la taille des objets les composant.

Comme le souligne Serra [110], il faut procéder en deux étapes bien distinctes. La première vise à transformer l'image originale : nous appliquons une granulométrie. La seconde vise à quantifier : nous estimons la distribution de taille. L'ultime étape consiste à segmenter l'image originale à partir des distributions de taille estimées.

#### 4.4.2 Granulométrie

Comme nous l'avons présenté précédemment, l'ouverture permet de filtrer les objets d'une image suivant un critère de taille lié à la taille de l'élément structurant utilisé pour faire cette analyse. L'analyse granulométrique, introduite par Matheron [89], consiste à itérer cette transformation (ou la transformation duale) à l'aide d'éléments structurants de taille croissante. A cette définition littérale de la granulométrie correspond une définition mathématique :

**Définition 4.4.1** Une granulométrie est une famille  $(\psi_\lambda)_{\lambda \in \mathbb{R}^+}$  d'opérateurs algébriques dépendant d'un unique paramètre  $\lambda$  qui vérifie les trois axiomes suivants [89] :

- (i)  $\forall \lambda \in \mathbb{R}^+, \psi_\lambda$  est croissante<sup>6</sup>.
- (ii)  $\forall \lambda \in \mathbb{R}^+, \psi_\lambda$  est anti-extensive<sup>7</sup>.
- (iii)  $\forall \lambda \in \mathbb{R}^+, \forall \mu \in \mathbb{R}^+, \psi_\lambda \circ \psi_\mu = \psi_\mu \circ \psi_\lambda = \psi_{\sup(\lambda, \mu)}$ .

En remarquant que la condition (iii) implique l'idempotence dans le cas où  $\lambda = \mu$ , on en déduit que  $\forall \lambda \in \mathbb{R}^+, (\psi_\lambda)$  est une ouverture au sens algébrique.

Dans [111], Serra donne une définition équivalente à celle donnée ci-dessus :

**Définition 4.4.2** Une granulométrie est une famille  $(\psi_\lambda)_{\lambda \in \mathbb{R}^+}$  d'opérateurs algébriques dépendant d'un unique paramètre  $\lambda$  qui vérifie les deux axiomes suivants :

- (i)  $\forall \lambda \in \mathbb{R}^+, \psi_\lambda$  est une ouverture algébrique.
- (ii)  $\forall \lambda \in \mathbb{R}^+, \forall \mu \in \mathbb{R}^+, \lambda \geq \mu, \psi_\lambda < \psi_\mu$ .

Dans la littérature, les granulométries les plus utilisées sont basées sur des ouvertures morphologiques par les homothétiques  $(\lambda B_1)_{\lambda \in \mathbb{R}^+}$  d'un élément structurant unité  $B_1$  convexe. Dans le cas d'une trame, i.e.  $B_1 \subset \mathbb{Z}^2$ , la granulométrie est paramétrisée par un entier  $r$  et  $rB_1 = \underbrace{B_1 \oplus B_1 \oplus \dots \oplus B_1}_{r \text{ fois}}$ .

Soulignons que certains définissent une anti-granulométrie ou granulométrie négative basée sur des fermetures morphologiques au lieu d'ouvertures.

### 4.4.3 Distribution de taille

Comme nous l'avons précisé précédemment, suite à une transformation morphologique, il y a une étape de quantification. C'est pourquoi, on associe souvent à une granulométrie une courbe, également appelée spectre de forme (*pattern spectrum*) [87], définie comme suit :

$$SF(\lambda) = 1 - \frac{Mes(\psi_\lambda(f))}{Mes(f)}$$

où  $Mes(\cdot)$  est le volume si  $f$  est une image en niveaux de gris et la surface si  $f$  est binaire.

Dans le cas d'une image discrète en niveaux de gris, le volume est défini dans une fenêtre comme la somme des niveaux de gris des pixels se situant dans cette fenêtre.

Notons que si l'élément structurant est de forme fixée,  $SF(\cdot)$  caractérise la taille des objets. On parlera alors de la distribution de taille.

A noter que la fonction  $SF(\cdot)$  peut être calculée localement sur une fenêtre centrée sur un pixel  $s$  [33] et sera dans ce cas notée  $SF_s(\cdot)$ .  $SF(\cdot)$  est une fonction croissante

<sup>6</sup> $\forall X, Y \subset \mathbb{R}^n, X \subset Y \Rightarrow \psi_\lambda(X) \subset \psi_\lambda(Y)$ .

<sup>7</sup> $\forall X \subset \mathbb{R}^n, \psi_\lambda(X) \subset X$ .

variant de 0 à 1 et représente une fonction de répartition. La dérivée discrète de  $SF(\cdot)$  est alors une fonction de densité de probabilité discrète notée  $dSF(\cdot)$  et définie comme suit :

$$dSF(\lambda) = \frac{Mes(\psi_\lambda(f)) - Mes(\psi_{\lambda+1}(f))}{Mes(f)}.$$

De même que pour  $SF(\cdot)$ ,  $dSF(\cdot)$  peut être définie localement et sera, dans ce cas, notée  $dSF_s(\cdot)$ .

Cette étape nous permet de quantifier des caractéristiques géométriques des objets comme la taille, ou encore la forme dans le cas où l'on utiliserait des éléments structurants de forme différente (e.g. linéaire).

#### 4.4.4 Application à l'analyse intra-urbaine

Nous avons appliqué l'approche décrite précédemment au problème particulier de l'analyse intra-urbaine. Rappelons que cette approche se décompose en deux phases distinctes. La première phase va consister à transformer les images satellitaires (4.8 (a) et (b)) des villes d'Aix en Provence et de Marseille par une série d'ouvertures par reconstruction de taille croissante. La seconde phase va consister à quantifier le paramètre de taille à partir des images précédemment obtenues.

##### 4.4.4.1 Granulométrie

Nous considérons une trame carrée dont l'élément structurant unité, plat, est l'approximation digitale d'un disque en 8 connexité (Figure 4.3). Cet élément, noté  $B_1$ , est non vide et convexe. Nous notons  $f$  l'image en niveaux de gris (4.8 (a) et (b)).

Nous choisissons comme opérateur de base de la granulométrie l'ouverture par reconstruction. La granulométrie est donc la famille d'opérateurs, paramétrée par l'entier  $r$  et définie de la manière suivante :

$$\psi_r(f) = \rho_f(f_{rB_1}), r \in \mathbb{N}, f \in \mathbb{R}^3. \quad (4.1)$$

$f \rightarrow \rho_f(f_{rB_1})$  est une ouverture algébrique, donc elle vérifie l'axiome (i) de la définition 4.4.2.

D'autre part, soit  $f \in \mathbb{R}^3$ , soient  $r$  et  $m$  appartenant à  $\mathbb{N}$  tel que  $r > m$ , alors par



(a): Aix en Provence ( $485 \times 506$ )



(b): Marseille ( $548 \times 598$ )

**Figure 4.8:** *Simulation SPOT5 à 5m*

définition  $f_{rB_1} \subset f_{mB_1}$ . Comme la reconstruction est une fonction croissante alors  $\forall f \in \mathbb{R}^3, \rho_f(f_{rB_1}) \subset \rho_f(f_{mB_1})$ . L'axiome (ii) de la définition 4.4.2 est vérifié.

Les résultats des premières itérations (i.e. pour des valeurs  $r \in \{1, 2, 3, 4, 5, 6\}$ ) de la granulométrie appliquée à l'image d'Aix en Provence sont donnés figure 4.9. En pratique, nous avons fait varier  $r$  de 1 à 34. En effet, à partir d'une certaine valeur de  $r$  l'image est complètement uniforme.

#### 4.4.4.2 Quantification

L'étape de quantification qui suit l'étape de transformation consiste, pour chaque pixel  $s$ , à estimer sur une fenêtre glissante centrée en  $s$ , les fonctions  $SF_s(\cdot)$  et  $dSF_s(\cdot)$  à partir de la série d'images ouvertes et reconstruites. Dans les exemples qui suivent, la taille de la fenêtre glissante a été déterminée de manière empirique de manière à résoudre le dilemme précision des contours et fiabilité des estimations.  $SF_s(\cdot)$  et  $dSF_s(\cdot)$  ont donc été estimées sur une fenêtre glissante de taille  $15 \times 15$ .

Notons que pour une valeur de  $r$  donnée, nous pouvons construire les images pour lesquelles la valeur du pixel  $s$  est soit  $SF_s(r)$ , soit  $dSF_s(r)$ . Des exemples de telles images sont présentés figure 4.10 pour différentes valeurs de  $r$ . Certaines zones des villes ressortent nettement comme par exemple le bâti dense sur la ville d'Aix en Provence (Figures 4.10 (a) et (b)) qui correspond au centre de la ville.

La méthode présentée ci-dessus nous a permis d'estimer en chaque pixel une fonction de distribution caractéristique de la distribution locale de la taille des objets. Cette information va être utilisée dans le modèle markovien défini ci-dessous dans le but de segmenter les images 4.8 (a) et (b) en plusieurs classes.

## 4.5 Modèle markovien de segmentation

### 4.5.1 Introduction

Après les étapes de transformation morphologique et de quantification, nous disposons pour chaque pixel  $s$  de l'image d'une fonction de répartition discrète  $SF_s(\cdot)$ .

Nous définissons donc un modèle de segmentation à partir des données extraites de l'image originale (i.e.  $SF_s(\cdot)$ ).

Le critère utilisé ici est le critère du MAP, déjà présenté dans 3.2.1.1.

Dans ce qui suit,  $X$  est le champ aléatoire relatif à l'image observée à valeurs dans  $\Omega^N$  ( $\Omega$  est l'ensemble des observations).

$L$  est le champ aléatoire relatif à l'image étiquetée à valeurs dans  $\Lambda^N$  ( $\Lambda$  est l'ensemble des étiquettes).

Les états font référence ici aux fonctions de répartitions estimées.

L'estimateur du MAP consiste donc à déterminer  $\hat{L}$  qui maximise :

$$P(X|L)P(L).$$

ou encore qui minimise l'énergie

$$\mathcal{U}(X, L) = \mathcal{U}_1(X, L) + \mathcal{U}_2(L) = -(\log(P(X|L)) + \log(P(L))).$$

Dans ce qui suit nous utilisons la représentation sous forme d'énergie.

## 4.5.2 Le modèle

Nos données étant des fonctions de répartition, nous avons choisi de définir le terme de vraisemblance du modèle à partir de la distance de Kolmogorov-Smirnov<sup>4</sup>  $d_{KS}(\cdot)$  entre la fonction de répartition  $SF_s(\cdot)$  associée au pixel  $s$  et la fonction de répartition  $SF_j(\cdot)$  associée à la classe  $j$ . Nous précisons dans le paragraphe traitant des résultats expérimentaux la manière dont sont estimées les fonctions  $SF_j(\cdot)$  associées aux différentes classes  $j$ . L'intérêt de cette méthode est qu'elle est non paramétrique. En effet, aucune hypothèse n'est faite sur les distributions (comme par exemple qu'elles soient normales). Par ailleurs, la distance  $d_{KS}(\cdot)$  est invariante par rapport aux transformations strictement monotones des données, i.e.  $d_{KS}(F, G) = d_{KS}(h(F), h(G))$ , où  $h$  est strictement croissante ou décroissante [42].

Le terme de vraisemblance du modèle est donc défini comme suit :

$$\mathcal{U}_1(X, L) = \sum_{s \in S} \left( \sum_{j \in \Lambda} \delta(l_s, j) d_{KS}(SF_s, SF_j) \right) \quad (4.2)$$

---

<sup>4</sup> $d_{KS}(F, G) = \max_t |F(t) - G(t)|$ , où  $F(\cdot)$  et  $G(\cdot)$  sont les fonctions de répartition associées à deux distributions.

Notons que dans [42, 65], les auteurs définissent l'énergie d'attache aux données de la manière suivante :

$$\mathcal{U}_1(X, L) = \sum_{s \in S} \left( \sum_{j \in \Lambda} \delta(l_s, j) \{2\Gamma(d_{KS}(SF_s, SF_j) > c_{KS}) - 1\} \right) \quad (4.3)$$

$\Gamma(d_{KS}(a, b) > c_{KS}) = 1$  si  $d_{KS}(a, b) > c_{KS}$  et  $\Gamma(d_{KS}(a, b) > c_{KS}) = 0$  sinon.

$c_{KS}$  est un paramètre qui est lié à la probabilité que la distance de Kolmogorov-Smirnov  $d_{KS}$  entre deux échantillons d'une même distribution soit plus grande que  $c_{KS}$ . Cette valeur peut être déterminée en utilisant des tables de statistique associées à la distribution de Kolmogorov limite.

Nous avons obtenu de meilleurs résultats en utilisant directement  $d_{KS}(SF_s, SF_j)$  dans le terme de vraisemblance au lieu d'utiliser le test statistique comme dans l'équation (4.3).

Le terme a priori est le modèle de Potts défini dans 3.2.1.3. L'énergie globale à minimiser est donc la suivante :

$$\mathcal{U}(X, L) = \sum_{s \in S} \left( \sum_{j \in \Lambda} \delta(l_s, j) d_{KS}(SF_s, SF_j) \right) - \beta \sum_{c=\{s,t\}, c \in \mathcal{C}} \delta(l_s, l_t) \quad (4.4)$$

Nous minimisons l'énergie ainsi définie en utilisant l'algorithme du recuit simulé décrit dans le paragraphe suivant.

### 4.5.3 Algorithme du recuit simulé

L'algorithme du recuit simulé est fondé sur l'analogie qui existe entre le processus physique de recuit notamment en métallurgie et la résolution de problèmes d'optimisation combinatoire.

Le principe est le suivant : le matériau est porté à une température élevée de manière à ce que les particules s'organisent pour aboutir à un état énergétique le plus faible possible. Puis, l'énergie est abaissée lentement de manière à ce que le matériau cristallise dans un état qui corresponde au minimum global d'énergie.

Ce principe a donc été transposé afin de résoudre des problèmes combinatoires. Nous avons utilisé l'algorithme du recuit avec dynamique de Métropolis [120]

(il en existe d'autres telle que la dynamique de Glauber [46], de Métropolis-Hasting [56], ou encore celle de Kawasaki [27]).

La minimisation de l'énergie  $\mathcal{U}(X, L)$  est donc effectuée avec l'algorithme du recuit simulé décrit ci-dessous :

#### Algorithme 4.5.1 (Algorithme du recuit simulé : dynamique de Métropolis)

- ① *Initialiser aléatoirement la configuration de départ et choisir une température initiale  $T_0$  suffisamment grande.*
- ② *Perturber la configuration précédente en un site.*
- ③ *Calculer  $\Delta\mathcal{U}$  la différence d'énergie entre l'ancienne configuration et la nouvelle.*  
*Si  $\Delta\mathcal{U} < 0$  alors on accepte la nouvelle configuration.*  
*Si  $\Delta\mathcal{U} \geq 0$ , on accepte la nouvelle configuration avec une probabilité égale à  $\exp\left(-\frac{\Delta\mathcal{U}}{T_k}\right)$ .*
- ④ *Continuer à l'étape 2 jusqu'à obtenir l'équilibre du système.*
- ⑤ *Diminuer la température selon la loi de décroissance choisie et reprendre à l'étape 2.*

En pratique, nous n'attendons pas l'équilibre du système à l'étape 4. L'algorithme est dit algorithme du recuit inhomogène [120].

Théoriquement pour que l'algorithme du recuit simulé converge vers le minimum global de l'énergie, il faut que la loi de température  $T_k$  décroisse moins rapidement que  $\frac{C_{ste}}{\ln k}$ , où  $k$  est le nombre d'itérations. En pratique, on est obligé de faire décroître la température plus rapidement. Nous avons choisi la loi de décroissance suivante :  $T_{k+1} = 0.98T_k$ . De ce fait, on perd la convergence vers la minimum global.

$T_0$  doit être choisi de manière à ce que toutes les transitions puissent avoir lieu avec une probabilité non nulle. Le critère d'arrêt de cet algorithme est un nombre de changements acceptés inférieur à 1% du nombre total de pixels dans l'image.

#### 4.5.4 Segmentation : résultats expérimentaux

Dans ce chapitre nous présentons les premiers résultats obtenus par la méthode décrite précédemment. Nous avons réalisé des tests sur les simulations SPOT5 à 5m

de résolution des villes d'Aix en Provence (Figure 4.8 (a)) et de Marseille (Figure 4.8 (b)).

Pour les tests expérimentaux, la segmentation est supervisée. Le nombre de classes ainsi que les paramètres associés à chacune d'elle sont donc déterminés par l'opérateur. Celui-ci sélectionne sur l'image originale une fenêtre qui contient une zone homogène (e.g. bâti dense, immeubles, pavillon). Sur cette fenêtre il estime la fonction de répartition  $SF_j(\cdot)$  associée à la classe  $j$  correspondante. La taille de la fenêtre doit être suffisamment grande pour obtenir des estimations fiables et suffisamment petite pour ne contenir qu'une seule texture. Précisons que le choix des classes a été réalisé à partir d'une observation visuelle de l'image à traiter. Des exemples de zones sélectionnées sont présentées figure 4.11. Sur chacune des deux images nous avons sélectionné 4 classes et les fonctions de répartition ont été estimées sur des fenêtres de taille  $51 \times 51$ .

L'algorithme du recuit simulé est initialisé avec une image d'étiquettes totalement aléatoire. La température initiale a été choisie égale à 10 afin de permettre, au départ, toutes les configurations. La règle de décroissance de la température et le critère d'arrêt sont précisés dans le paragraphe présentant l'algorithme du recuit simulé.

Les résultats de la segmentation des deux villes d'Aix en Provence et de Marseille en 4 classes sont présentés figures 4.12 (a) et (b).

Sur l'image d'Aix en Provence (Figure 4.12 (a)) le bâti dense (qui correspond pour l'essentiel au centre ville) est bien détecté. Il apparaît nettement sur l'image de paramètre (Figure 4.10 (a) et (b)). Cette classe est celle qui est la mieux détectée par notre approche. Notons que dans [47], l'auteur qui traite la même image obtient également un taux de détection nettement plus important pour cette classe. Les classes qui correspondent aux pavillons et aux immeubles sont moins bien détectées. De nombreuses confusions sont commises entre ces deux classes. Cette mauvaise détection peut être due à la fois au choix des fenêtres d'estimation mais également à l'analyse de la texture qui ne prend en compte que la taille des objets et non leur forme. Il aurait été intéressant d'essayer de faire varier la forme des éléments structurants lors de l'analyse granulométrique afin de voir l'apport d'une telle information pour distinguer les différentes classes.

Sur l'image de Marseille (Figure 4.12 (b)), les classes pavillons et immeubles sont mieux séparées que sur l'image d'Aix en Provence. La mer est également très bien détectée. Une analyse plus fine des résultats par un thématicien ayant

Nom Image Taille	Capteur Résolution	CPU granulométrie	CPU quantification	CPU recuit simulé
Aix 485×506	SPOT5 5m	17min18s	1min5s	14min56s
Marseille 548×598	SPOT5 5m	25min50s	1min29s	21min44s

**Tableau 4.3:** Valeurs des paramètres et temps de calcul sur station SUN ultra1 (124MB, 167 MHz).

une bonne connaissance du terrain s'avère nécessaire à la fois pour caractériser la méthode mais également pour la réorienter.

Notons par ailleurs que texture correspondant à la classe immeubles sur la ville d'Aix en Provence est visuellement très différente de celle de la ville Marseille. Les bâtiments n'ont pas la même forme, ni la même taille. C'est pourquoi la mise en œuvre d'une méthode d'analyse très générale des zones urbaines semble difficile à réaliser et nécessite au moins une adaptation aux divers types de tissu urbain que l'on peut rencontrer.

Il nous paraît nécessaire de souligner que le choix des classes et l'estimation des paramètres ont été réalisés de manière supervisée. Les différents tests expérimentaux ont montré que la classification finale dépendait fortement de ce choix et de l'estimation des paramètres de ces classes. Sur ce type de problème, il serait nécessaire qu'un géographe ayant une bonne connaissance du terrain ou qu'un photo-interprète opère la phase supervisée. D'autre part, l'estimation des paramètres associés aux diverses classes a été réalisée sur des fenêtres de taille  $51 \times 51$ . Un photo-interprète pourrait lui-même à partir de sa connaissance du terrain délimiter des zones d'apprentissage plus fiables et de forme quelconque. Enfin, notons que la granulométrie basée sur une série d'ouvertures par reconstruction est une opération coûteuse en temps de calcul (Tableau 4.3) dans la mesure où chaque ouverture par un élément structurant d'une certaine taille est suivie d'une reconstruction.

## 4.6 Conclusion

Nous avons proposé une approche fondée sur la morphologie mathématique. L'analyse de l'image originale est réalisée en deux étapes distinctes. La première, appelée granulométrie, a consisté à transformer l'image en niveaux de gris par une série d'ouvertures par reconstruction avec un élément structurant plat de taille croissante. La seconde étape, appelée quantification, a consisté à extraire pour chaque pixel une fonction de répartition caractéristique de la distribution locale de la taille des objets.

L'originalité du travail a été d'utiliser cette information dans un modèle markovien de segmentation dont le terme de vraisemblance est défini à partir de la distance de Kolmogorov-Smirnov entre deux distributions.

L'utilisation d'une granulométrie, très peu utilisée dans le cadre d'analyse d'images satellitales, fondée sur des ouvertures par reconstruction est très coûteuse en temps de calcul. En effet, chaque ouverture par un élément structurant de taille donnée est suivie d'une reconstruction.

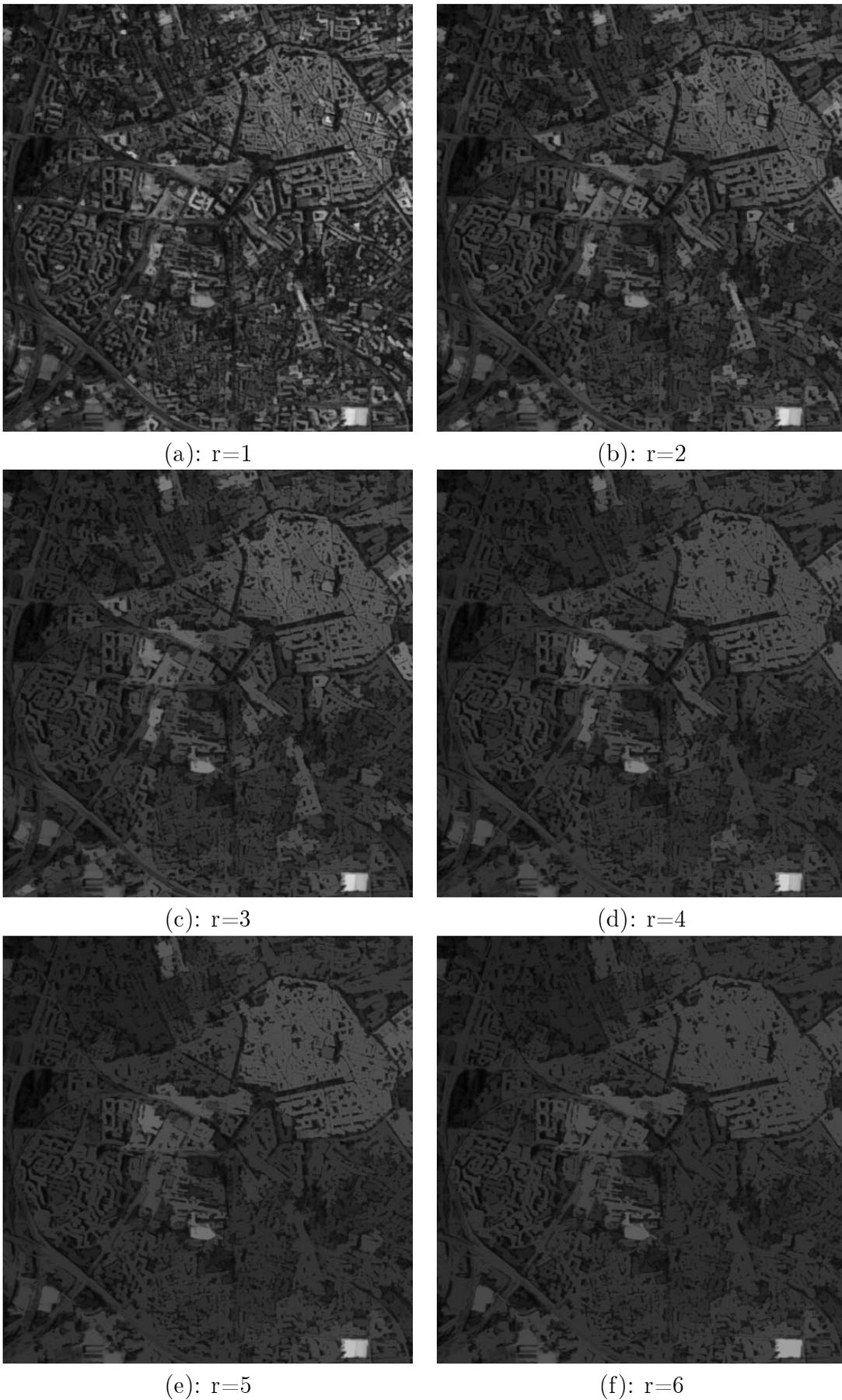
Les résultats expérimentaux obtenus par la méthode proposée comportent de nombreuses confusions notamment entre les classes pavillons et immeubles. Le bâti dense est quant à lui bien délimité. La méthode proposée est donc une première approche pour l'analyse intra-urbaine mais demande très certainement à être améliorée. Soulignons par ailleurs que la phase du choix des classes et d'estimation des paramètres dont dépend le résultat final nécessite l'intervention d'un opérateur qui a une bonne connaissance du terrain.

Enfin, sur les images SPOT5 à 5m les structures que l'on distingue à l'oeil ne sont pas à la même échelle. En effet, le vieux centre ville d'Aix en Provence apparaît comme un seul et même bloc, i.e. la résolution est trop faible pour pouvoir distinguer du bâti séparé par des rues très étroites. Par ailleurs, les immeubles hauts se distinguent nettement les uns des autres. Ainsi, dans un cas l'objet est l'immeuble lui-même dans l'autre cas c'est le quartier. Cette différence nous paraît être un désavantage de cette résolution.

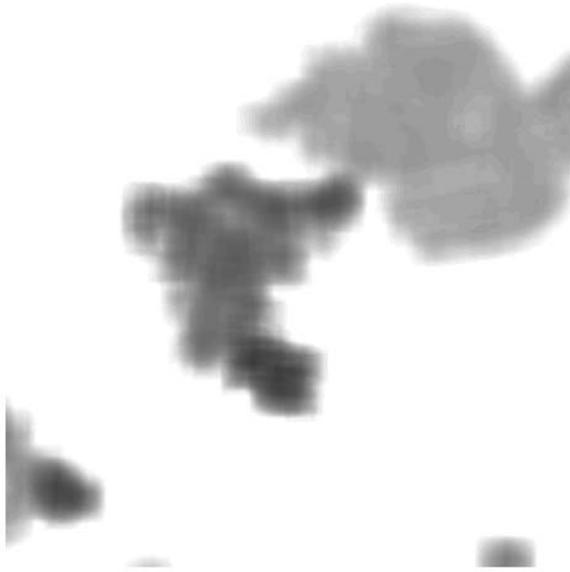
Par ailleurs, on peut se poser la question de l'intérêt d'une telle résolution, pour ce qui est de l'analyse urbaine, alors même que d'autres pays vont proposer des satellites civils à une résolution métrique qui semble mieux appropriée à une étude précise de la complexité du paysage urbain. Ainsi, suite à l'échec du lancement du satellite *Earlybird* en décembre 1997, la société *Earthwatch Consortium* a choisi de passer directement au programme suivant, i.e. au lancement du satellite

*Quickbird* ayant une résolution de 1m.

En effet, dans la plupart des applications visées la résolution nécessaire est soit une résolution submétrique (télécommunications, cartographie BD Topo® de l'IGN), soit une résolution décamétrique (BD Carto® de l'IGN, suivi de population dans les zones tropicales).



**Figure 4.9:** Exemple de granulométrie sur l'image (Figure 4.6 (a)), premières itérations :  $r \in \{1, 2, 3, 4, 5, 6\}$ .



(a): Pour  $r = 8$ ,  $dSF_s(r)$



(b): Pour  $r = 9$ ,  $dSF_s(r)$



(c): Pour  $r = 3$ ,  $dSF_s(r)$



(d): Pour  $r = 4$ ,  $dSF_s(r)$

**Figure 4.10:** (a)(b): *Aix en Provence*, (c)(d): *Marseille*

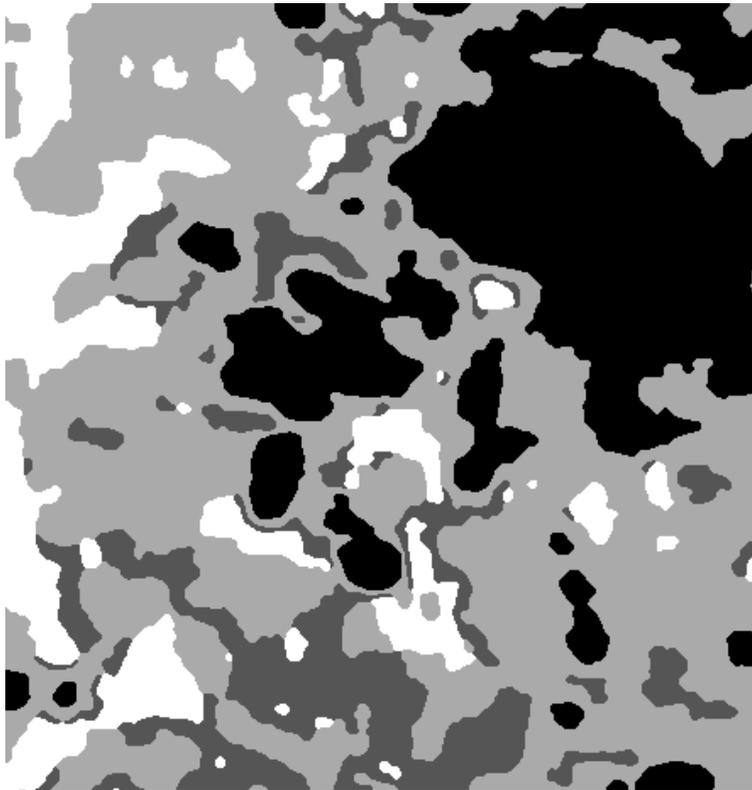


(a): Aix en Provence

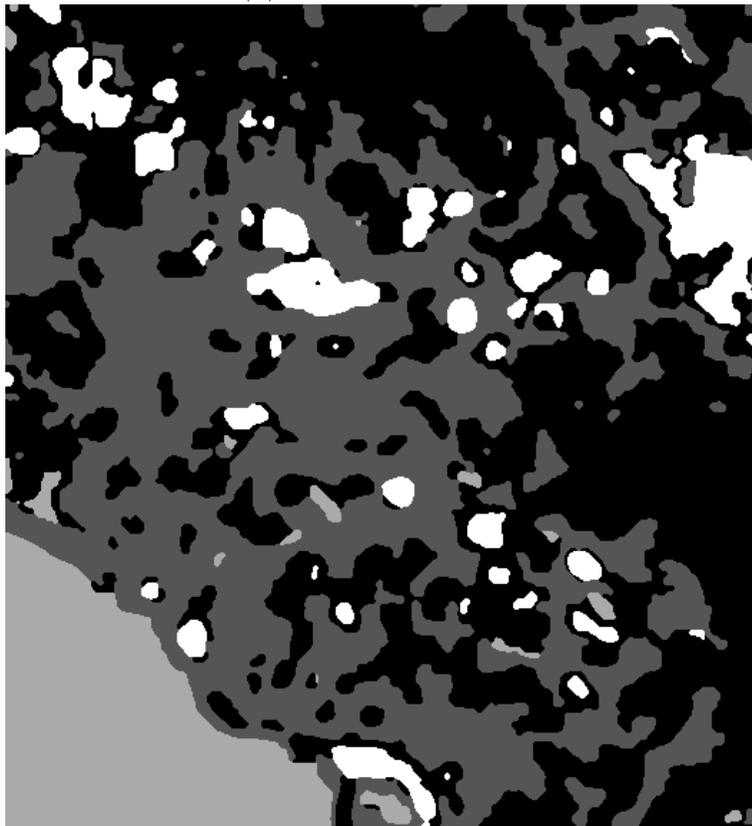


(b): Marseille

**Figure 4.11:** *Classes sélectionnées*



(a): Aix en Provence



(b): Marseille

**Figure 4.12:** *Segmentation finale en 4 classes.*

## Conclusion

Après une synthèse du travail effectué, nous donnons quelques axes possibles en vue d'améliorer, de réorienter les futures recherches à la fois sur les problèmes généraux d'analyse de texture, d'estimation des paramètres, de classification, de segmentation et sur les applications plus particulières d'analyse du milieu l'urbain.

Nous avons proposé de développer des méthodes de traitement d'image afin d'analyser les images satellitales en milieu urbain. Dans cette optique, nous avons défini un nouveau paramètre de texture. Ce paramètre est estimé suite à la modélisation du champ de luminance par 8 modèles markoviens basés sur des chaînes.

Sur l'aspect de l'analyse de la texture, l'originalité vient du choix du système de voisinage par rapport aux systèmes classiques 4 ou 8 connexes. Nous avons également proposé une approche originale de normalisation des paramètres dans les différentes directions fondée sur une méthode de renormalisation par décimation issue de la physique statistique. Il faut quand même souligner que cette étape complexifie l'analyse de la texture. Des comparaisons avec des approches classiques ont mis en évidence l'amélioration des résultats d'un point de vue qualitatif.

Le choix du paramètre de texture, comme combinaison linéaire des valeurs centrales estimées en chaque pixel, a été réalisé à partir de l'étude sur plusieurs zones homogènes extraites d'images à différentes résolutions, provenant de différents capteurs. Il aurait été intéressant d'utiliser une approche moins empirique. En effet, il existe des méthodes [30, 41] permettant de sélectionner parmi un jeu de paramètres, ceux permettant d'obtenir la meilleure discrimination entre les classes. Cette sélection doit permettre de diminuer les temps de traitement des procédures ultérieures (e.g. classification, segmentation). Les paramètres sélectionnés doivent alors permettre d'obtenir un résultat précis.

La méthode de classification floue proposée est intéressante pour deux raisons. Elle permet tout d'abord d'estimer le nombre de classes et les paramètres associés. Elle fournit par ailleurs une information supplémentaire par rapport à une méthode de classification "dure", i.e. la probabilité d'appartenance de chaque pixel à chacune des classes. Ces informations sont finalement utilisées dans un modèle markovien de segmentation classique. Cette méthode est d'ailleurs applicable à bien d'autres domaines, i.e. tous ceux qui nécessitent une classification/segmentation d'image.

Finalement, les résultats obtenus sont satisfaisants à plusieurs points de vue. Premièrement, les contours des zones urbaines sont précis et francs. D'autre part, l'analyse proposée permet d'extraire le masque urbain et également de caractériser dans l'urbain les zones végétales de surface suffisante. Enfin, de nombreux tests sur des données réelles ont montré que l'approche proposée était robuste vis à vis de la résolution des images et également vis à vis du type de capteur. Il peut être intéressant d'avoir une approche générale qui permet de travailler à la fois sur de l'optique et du radar.

Pour les futures recherches, il nous paraît intéressant d'envisager de grouper

la phase de classification et de segmentation dans un même critère [26]. La méthode envisagée consiste à associer à chaque pixel un vecteur de paramètres dont chaque composante correspond au degré d'appartenance du pixel à une des classes au lieu de lui associer une seule étiquette. Le critère global serait alors la somme du critère entropique que nous avons proposé et d'un terme de régularisation, de type Potts.

La méthode d'analyse intra-urbaine utilise une approche issue de la morphologie mathématique : la granulométrie. Largement utilisée dans des domaines tels que l'analyse des structures des roches, l'imagerie médicale, elle l'est moins dans le domaine de la télédétection. L'approche considérée permet de faire une analyse locale de la distribution de la taille des objets.

La méthode proposée, si elle a permis d'extraire le bâti dense, n'a pas permis d'extraire de manière satisfaisante les zones d'immeubles et les zones de pavillons. Des confusions ont été faites sur ces deux classes. Il s'agit d'une première approche d'analyse intra-urbaine qui nécessite très certainement des améliorations à la fois sur la détermination des classes et l'estimation des paramètres associés, que sur la méthode elle-même. En effet, nous nous sommes limités à une seule forme d'élément structurant. Il serait intéressant d'envisager une étude à partir d'une série d'éléments structurants de forme variable afin d'intégrer dans l'analyse à la fois la taille et la forme des objets.

Un des problèmes majeurs auquel nous avons été confronté est la validation des résultats. En effet, elle a été faite d'un point de vue qualitatif avec pour certains résultats une validation par un thématicien et pour d'autres par comparaison avec des données aériennes. Cependant, nous n'avions à notre disposition aucune vérité terrain. Par ailleurs, la ville reste une entité dont les contours sont imprécis. La seule définition "mathématique" de la ville est la ZMU (Zone Morphologique Urbaine) utilisée notamment dans le cadre du projet EUROSTAT. Dans ce cas, la validation des classifications se fait à partir de données cadastrales, ou d'autres données exogènes par un photo-interprète. Les validations qui sont parfois faites par des gens issus du traitement d'image et qui consistent à calculer une simple matrice de confusion à partir d'une vérité terrain plus ou moins grossière et d'affirmer par la suite obtenir une classification correcte à  $n$  % est très discutable, tant au niveau du chiffre donné ( $n$  %) que de la méthode employée. Ceci d'autant plus dans le domaine de l'urbain où les limites des zones sont beaucoup moins précises que dans le domaine agricole (limite de parcelles).

A ce sujet le lecteur pourra se reporter à l'article suivant [75] qui propose une méthode de mesure de la qualité des résultats qui tient compte des aspects de géométrie, de sémantique, d'exhaustivité. Cette approche tient compte de types

d'erreurs différents et permet d'obtenir une quantification de chacune d'elle. L'article met bien en évidence toute la difficulté de la validation des résultats. En effet, la méthode, tout en étant relativement complète, ne permet d'obtenir selon les auteurs qu'une idée du niveau de qualité.

Notre application reste très générale : extraire et analyser l'urbain. Il n'y a pas d'utilisateur précis, d'interlocuteur privilégié comme cela est le cas lors de projets tels ALTLAS, EUROSTAT. Or il est à mon sens inévitable si l'on veut améliorer l'automatisation du processus d'analyse de l'urbain et la qualité du résultat d'avoir un interlocuteur précis, et une application bien identifiée. En effet, la qualité, la précision du résultat recherché dépend fortement de l'utilisation qui doit être faite de cette analyse. Les critères de qualité peuvent être une simple validation visuelle, une segmentation finale proche d'une segmentation manuelle de la même zone, une bonne détection, ou bien encore une localisation très précise, etc. En outre, pour certaines applications les utilisateurs préféreront une surdétection des zones bâties; pour d'autres, en revanche, il est nécessaire que les zones détectées soit très précises et préféreront une sous détection.

Enfin, je pense que si l'on disposait de données exogènes (e.g SIG) beaucoup d'ambiguïtés pourraient être levées. Ainsi, la méthodologie utilisée par certains géographes [76] me paraît très intéressante. Elle consiste à diviser l'image à segmenter en zones géographiques homogènes (e.g. plaines basses) à partir de la connaissance a priori que le géographe a du terrain. Cette analyse préliminaire permet notamment d'éviter de nombreuses confusions. Ainsi, dans le cas d'objets ayant des textures très proches (e.g. quartiers neufs et plantations de cellulose), le fait de tenir compte de la région à laquelle ils appartiennent permet de lever l'ambiguïté sur leur nature.

En effet, l'utilisation de l'image comme seule source de données a ses limites. La connaissance du terrain, ou l'apport de données exogènes permettrait déjà d'orienter la phase d'extraction des zones urbaines et d'éviter probablement nombre de confusions. L'interprétation totalement automatique d'images satellitales atteint ici à mon sens ses limites. L'approche proposée pourrait constituer une aide à la photo-interprétation.

En conclusion, il me semble absolument nécessaire si l'on veut améliorer les résultats que le traiteur d'image travaille davantage en collaboration avec l'utilisateur potentiel de son travail, e.g. le géographe.

## Renormalisation par décimation

Le but est ici de calculer les paramètres du réseau décimé en fonction des paramètres du réseau complet (Figure 2.7).

Pour cela, il convient d'intégrer la loi jointe du modèle par rapport aux sites que l'on souhaite voir disparaître.

Les calculs sont effectués ici en décimant un site sur deux. Ceci est un cas particulier. En effet, pour traiter le problème présenté dans le chapitre 1.2.3 il convient de décimer davantage le réseau et d'intégrer sur 11, 16 et 27 variables.

Dans ce qui suit  $E[.]$  est la fonction partie entière.

La loi jointe est la suivante :

$$P(x_1, x_2, \dots, x_N) \propto \exp \left( -\beta \left( \sum_{k=1}^N \lambda(x_k - \mu)^2 + \sum_{k=1}^{N-1} (x_k - x_{k+1})^2 \right) \right)$$

Nous supposons par ailleurs des conditions aux bords toroïdales. Dès lors,  $x_1$  a deux voisins  $x_0$  et  $x_2$ , avec  $x_0 = x_N$ . De même,  $x_N$  a deux voisins  $x_{N-1}$  et  $x_1$ .

$$P(x_2, x_4, \dots, x_{2E[\frac{N}{2}]}) \propto \int_{x_1, x_3, \dots, x_{2E[\frac{N-1}{2}]+1}} \exp \left( -\beta \left( \sum_{k=1}^N \lambda(x_k - \mu)^2 + \sum_{k=1}^{N-1} (x_k - x_{k+1})^2 \right) \right) dx_1 \dots$$

$$P(x_2, x_4, \dots, x_{2E[\frac{N}{2}]}) \propto \exp \left( -\beta \sum_{k=1}^{E[\frac{N}{2}]} \lambda(x_{2k} - \mu)^2 \right) * I$$

Avec :

$$I = \prod_{k=0}^{E[\frac{N-1}{2}]} \left( \int_{x_{2k+1}} \exp \left( -\beta (\lambda(x_{2k+1} - \mu)^2 + (x_{2k+1} - x_{2k})^2 + (x_{2k+1} - x_{2k+2})^2) \right) dx_{2k+1} \right)$$

On pose :

$$I_{x_{2k+1}} = \int_{x_{2k+1}} \exp(-\beta(\lambda(x_{2k+1} - \mu)^2 + (x_{2k+1} - x_{2k})^2 + (x_{2k+1} - x_{2k+2})^2)) dx_{2k+1}$$

Le calcul de  $I_{x_{2k+1}}$  nous donne :

$$I_{x_{2k+1}} = \sqrt{\frac{\pi}{\beta(2 + \lambda)}} \exp\left(-\frac{\beta}{2 + \lambda} ((x_{2k+2} - x_{2k})^2 + \lambda(x_{2k+2} - \mu)^2 + \lambda(x_{2k} - \mu)^2)\right)$$

Nous rappelons que  $x_0$  et  $x_{N+1}$  sont définis par les conditions aux bords précisées ci-dessus.

On en déduit donc que :

$$P(x_2, x_4, \dots) \propto \exp\left(-\frac{\beta}{2 + \lambda} \left(\sum_{k=1}^{E[\frac{N-2}{2}]} (x_{2k} - x_{2k+2})^2 + \sum_{k=1}^{E[\frac{N}{2}]} (\lambda(2 + \lambda) + 2\lambda)(x_{2k} - \mu)^2\right)\right)$$

$$P(x_2, x_4, \dots) \propto \exp\left(-\beta_1 \left(\sum_{k=1}^{E[\frac{N-2}{2}]} (x_{2k} - x_{2k+2})^2 + \sum_{k=1}^{E[\frac{N}{2}]} \lambda_1 (x_{2k} - \mu)^2\right)\right)$$

On obtient donc les 2 relations suivantes :

$$\beta_1 = \frac{\beta}{2 + \lambda}$$

$$\lambda_1 = \lambda(4 + \lambda)$$

Un calcul analogue permet d'obtenir les résultats en décimant davantage le réseau et en intégrant sur 11, 16 et 27 variables . Ces calculs ont été effectués avec le logiciel Maple.

# Bibliographie

- [1] Ahuja N. et Rosenfeld A. – Mosaic models for textures. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 3, no. 1, janvier 1981, pp. 1–11.
- [2] Akaike H. – A new look at the statistical model identification. *IEEE Trans. on automatic control*, vol. 19, no. 6, décembre 1974, pp. 716–723.
- [3] Alata O., Baylou P. et Najim M. – Texture characterization based on 2D reflection coefficients. *In: ICASSP*, pp. 2539–2542. – Detroit, 1995.
- [4] Armand M. – Cahier A4 télédétection, urbanisme et aménagement. – Les cahiers pédagogiques du GDTA, 1995.
- [5] Armand M. – Les données SPOT: quels apports à l’urbanisme. *Bul. S.F.P.T.*, no. 139, mars 1995, pp. 32–38.
- [6] Azencott R. – Image analysis and Markov fields. *In: ICIAM*, pp. 53–61. – Paris, 1987.
- [7] Banerji A. et Goutsias J. – A morphological approach to automatic mine detection problems. *IEEE trans. on aerospace and electronic systems*, vol. 34, no. 4, octobre 1998, pp. 1085–1096.
- [8] Baraldi A. et Parmiggiani F. – Urban area classification by multispectral SPOT images. *IEEE Trans. on geoscience and remote sensing*, vol. 28, no. 4, juillet 1990, pp. 674–680.
- [9] Bardinnet C., Le Tourneau F. M. et Jalon J. – Etude par télédétection du problème urbain dans la vallée de Kathmandu. *In: High Mountain Remote Sensing*. – Mandoza Argentine, 1994.
- [10] Batty M. et Longley P. – *Fractal Cities*. – Academic Press, 1994.

- [11] Baudoin A., Fratter C. et Munier P. – Le programme SPOT 5 et la préparation des utilisations. *Bul. S.F.P.T.*, no. 141, janvier 1996, pp. 9–21.
- [12] Benharrosh J-M. – *Extraction de thèmes cartographiques dans les images satellitaires ou aériennes. Application à la génération de quick-looks adaptatifs et à la compression d'images.* – Thèse de Doctorat, UNSA, avril 1998.
- [13] Besag J. – On the statistical analysis of dirty pictures. *Journal of the royal statistical society*, no. 48, 1986, pp. 259–302. – Série B.
- [14] Besag J. E. – Spatial interaction and the statistical analysis of lattice system. *Journal of the royal statistical society, Series B*, no. 36, 1974, pp. 192–236.
- [15] Bezdek J.C. – *Pattern recognition with fuzzy objective function algorithms.* – Plenum Press, New-York, 1981.
- [16] Bouman C. et Liu B. – Multiple resolution segmentation of textured images. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 13, no. 2, février 1991, pp. 99–113.
- [17] Bovik A. C. – Analysis of multichannel narrow-band filters for image texture segmentation. *IEEE Trans. on signal processing*, vol. 39, no. 9, septembre 1991, pp. 2025–2043.
- [18] Cardy J. – *Scaling and Renormalization in Statistical Physics.* – Cambridge University Press, 1996.
- [19] Chatterjee S. – *Markov Random Fields: Theory and Applications*, chap. Classification of Natural Textures Using Gaussian Markov Random Fields Models, pp. 159–177. – Academic Press, 1993.
- [20] Chellappa R. et Kashyap R. – Texture synthesis using 2D non-causal autoregressive models. *IEEE Trans. on acoustics, speech, and signal processing*, vol. 33, no. 1, février 1985, pp. 194–203.
- [21] Cohen F. S. et Cooper D. B. – Simple parallel hierarchical and relaxation algorithms for segmenting non causal Markovian random fields. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 9, no. 2, mars 1987, pp. 195–219.
- [22] Connors R. W., Trivedi M. M. et Harlow C. A. – Segmentation of high-resolution urban scene using texture operators. *CVGIP*, vol. 25, 1984, pp. 273–310.

- [23] Crimmins T. R. – Geometric filter for speckle reduction. *Applied optics*, vol. 24, no. 10, mai 1985, pp. 1438–1443.
- [24] Cross G. R. et Jain A. K. – Markov random field texture models. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 5, no. 1, janvier 1983, pp. 25–39.
- [25] Descombes X. – *Champs markoviens en analyse d'images*. – Thèse de Doctorat, ENST (93E026), décembre 1993.
- [26] Descombes X. et Kruggel F. – A Markov pixon information approach for low level image description. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 21, no. 6, juin 1999, pp. 482–494.
- [27] Descombes X. et Pechersky E. – Metropolis vs Kawasaki dynamic for image segmentation based on Gibbs models. In: *EMMCVPR*. – York, juillet 1999.
- [28] Descombes X. et Prêteux F. – Topology and parameter estimation in mrf modeling. In: *SPIE, Neural and Stochastic Methods in Image and Signal Processing II*, pp. 156–166. – San Diego, juillet 1993.
- [29] Descombes X., Sigelle M. et Prêteux F. – Estimating Gaussian Markov random field parameters in a non-stationary framework: Application to remote sensing imaging. *IEEE Trans. on image processing*, vol. 8, no. 4, avril 1999, pp. 490–503.
- [30] Devijver P. et Kittler J. – *Pattern recognition: a statistical approach*. – Prentice-Hall, 1982.
- [31] Dougherty E. R., Newell J. T. et Pelz J. B. – Morphological texture-based maximum-likelihood pixel classification based on local granulometric moments. *Pattern recognition*, vol. 25, no. 10, 1992, pp. 1181–1198.
- [32] Dougherty E. R. et Pelz J. B. – Morphological granulometric analysis of electrophotographic images-size distribution statistics for process control. *Optical engineering*, vol. 30, no. 4, avril 1991, pp. 438–445.
- [33] Dougherty E. R., Pelz J. B., Sand F. et Lent A. – Morphological image segmentation by local granulometric size distributions. *Journal of electronic imaging*, vol. 1, no. 1, janvier 1992, pp. 46–60.
- [34] Dubes R. C. et Jain A. K. – Random fields models in image analysis. *Journal of applied statistics*, vol. 16, no. 2, 1989, pp. 131–164.

- [35] Elfadel I. M. et Picard R. W. – Gibbs random fields, co-occurrences, and texture modelling. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 16, no. 1, janvier 1993, pp. 24–37.
- [36] Fogel I. et Sagi D. – Gabor filters as texture discriminator. *Biological cybernetics*, vol. 61, 1989, pp. 103–113.
- [37] Forbes F. et Raftery A. R. – Bayesian morphology: Fast unsupervised bayesian image analysis. *Journal of the American Statistical Association - Theory and Methods*, vol. 94, no. 446, juin 1999, pp. 555–568.
- [38] Francos J. M., Meiri A. Z. et Porat B. – A unified texture model based on a 2D Wold-like decomposition. *IEEE trans. on signal processing*, vol. 41, no. 8, août 1993, pp. 2665–2678.
- [39] Frigui H. et Krishnapuram R. – A robust clustering algorithm based on competitive agglomeration and soft rejection of outliers. *In: CVPR*, pp. 550–555. – San Francisco, juin 1996.
- [40] Fu K. S. – *Syntactic Pattern Recognition and Applications*, chap. 12- A Syntactic Approach to Texture Analysis, pp. 445–480. – Prentice Hall, 1982.
- [41] Fukunaga K. – *Introduction to statistical pattern recognition*. – Academic Press, 1990. Seconde édition.
- [42] Geman D., Geman S., Graffigne C. et Dong P. – Boundary detection by constrained optimization. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 12, no. 7, juillet 1990, pp. 609–628.
- [43] Geman S. et Geman D. – Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 6, no. 6, novembre 1984, pp. 721–741.
- [44] Gidas B. – A renormalization group approach to image processing problems. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 11, no. 2, février 1989, pp. 164–180.
- [45] Gimel'farb G. L. – Gibbs models for bayesian simulation and segmentation of piecewise-uniform textures. *In: ICPR*, pp. 760–764. – Vienne, août 1996.
- [46] Glauber R. J. – Coherent and incoherent states of the radiation field. *Phys. Rev.*, no. 131, 1963, pp. 2766–2788.

- [47] Gouinaud C. – *Traitement d'images satellitaires pour la détection d'agglomérations*. – Thèse de Doctorat, ENST, 1996.
- [48] Gouinaud C., Tupin F. et Maitre H. – Potential and use of radar images for characterization and detection of urban areas. *In: IGARSS*, pp. 474–476. – Lincoln, mai 1996.
- [49] Gratin C., Vitrià J., Moreso F. et Serón D. – Texture classification using neural networks and local granulometries. *In: International Conference on Mathematical morphology and its applications to image processing*, éd. par Serra J. et Soille P. pp. 309–316. – Kluwer academic publishers.
- [50] Greenspan H., Belongie S., Goodman R. et Perona P. – Rotation invariant texture recognition using a steerable pyramid. *In: ICPR*, pp. 162–167. – Jérusalem, octobre 1994.
- [51] Gull S. F. et Skilling J. – Maximum entropy method in image processing. *Proceedings of the institution of electrical engineers*, vol. 131, 1984, pp. 646–659.
- [52] Haas A., Matheron G. et Serra J. – *Morphologie mathématique et granulométries en place*, pp. 735–753. – 1967 -vol. XI. 1ère partie.
- [53] Haas A., Matheron G. et Serra J. – *Morphologie mathématique et granulométries en place*, pp. 767–782. – 1967 -vol. XII. 2ème partie.
- [54] Haley G. M. et Manjunath B. S. – Rotation-invariant texture classification using modified Gabor filters. *In: ICIP*, pp. 262–265. – Washington, septembre 1995.
- [55] Haralick R. M. – Statistical and structural approaches to texture. *Proc. of the IEEE*, vol. 67, no. 5, mai 1979, pp. 786–804.
- [56] Hastings W. K. – Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, no. 57, 1970, pp. 97–109.
- [57] Heijmans H. – Theoretical aspects of grey level morphology. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 13, no. 6, 1991, pp. 568–592.
- [58] Hoffman R. et Jain A. K. – Segmentation and classification of range images. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 9, no. 5, 1987, pp. 608–620.

- [59] Houzelle S. et Giraudon G. – Data fusion using SPOT and SAR images for bridge and urban area extraction. *In: IGARSS*, pp. 1455–1458. – Helsinki, juin 1991.
- [60] Huang K. – *Quantum Field Theory: From Operators to Path Integrals*. – John Wiley, New York, 1998.
- [61] Hubel D.H. – *Eye, brain, and vision*. – Freeman, 1988.
- [62] Huet C., Tonon M., Le Men H. et Jamet O. – Towards an automatic updating process for landcover cartography using Spot images interpretation. *In: European Symposium on satellite remote sensing III*. – Taormine, septembre 1996.
- [63] Jain A. K. et Farrokhnia F. – Unsupervised texture segmentation using Gabor filters. *Pattern recognition*, vol. 24, no. 12, 1991, pp. 1167–1186.
- [64] Julesz B. et Bergen J. R. – *Readings in Computer Vision, Issues, Problems, Principles and Paradigms*, chap. Textons, The Fundamental Elements in Preattentive Vision and Perception of Textures, pp. 243–256. – Morgan Kaufmann, 1987.
- [65] Kervrann C. et Heitz F. – A Markov random field model-based approach to unsupervised texture segmentation using local and global statistic. *IEEE Trans. on image processing*, vol. 4, no. 6, juin 1995, pp. 856–862.
- [66] Khinchin A. I. – *Mathematical foundation of information theory*, 1957. Dover.
- [67] Klein J. C. et Serra J. – The texture analyser. *Journal of microscopy*, vol. 95, no. 2, 1972, pp. 343–356.
- [68] Krishnamachari S. et Chellappa R. – GMRF models and wavelet decomposition for texture segmentation. *In: ICIP*, pp. 568–571. – Washington, octobre 1995.
- [69] Krishnamachari S. et Chellappa R. – Multiresolution GMRF models for texture segmentation. *In: ICASSP*, pp. 2407–2410. – Détroit, mai 1995.
- [70] Krishnamurthy R., Woods J. W. et Francos J. M. – Adaptive restoration of textured images with mixed spectra. *IEEE Trans. on image processing*, vol. 5, no. 4, avril 1996, pp. 648–652.

- [71] Krishnapuram R. et Keller J. M. – *Fuzzy and possibilistic clustering methods for computer vision*. – 1994, *SPIE institute series, in neural and fuzzy systems*, vol. 12. pp. 133-159.
- [72] Lantuéjoul C. et Maisonneuve F. – Geodesic methods in image analysis. *Pattern recognition*, vol. 17, 1984, pp. 117–187.
- [73] Laws K. – *Texture Image Segmentation*. – Thèse de Doctorat, University of Southern California, janvier 1980.
- [74] Le Men H. – Cartographie semi-automatique de l'évolution de l'occupation des sols par télédétection. – Habilitation à diriger des recherches, décembre 1996.
- [75] Le Men H. et Jamet O. – Qualité de processus d'interprétation et qualité des résultats: un exemple en cartographie d'occupation du sol. *CFC*, no. 142/143, déc 94/mars 95, pp. 182–192.
- [76] Le Tourneau F.-M., Lorette A., Rudant J-P. et Descombes X. – Utilisation d'une séquence d'images ERS et d'une méthode d'extraction automatique des zones urbaines pour le suivi de l'extension de la ville de Macap (Amap , Brésil). *soumis au bulletin de la SFPT*, juin 1999.
- [77] Lemonnier H. – Apport de la précision spatiale des futures données SPOT 5 dans le contexte du projet EUROSTAT sur la délimitation des agglomérations. *Bul. S.F.P.T.*, no. 141, janvier 1996, pp. 80–83.
- [78] Levy-vehel J., Mignot P. et Berroir J-P. – Texture and multifractals: new tools for image analysis. – INRIA research report RR-1706, juin 1992.
- [79] Li S. Z. – *Markov Random Field Modeling in Computer Vision*. – Springer-Verlag, 1995.
- [80] Liang Z., MacFall J. R. et Harrington D. P. – Parameter estimation and tissue segmentation from multispectral MR images. *IEEE trans. on medical imaging*, vol. 13, no. 3, septembre 1994, pp. 441–449.
- [81] Liu F. et Picard R. W. – Periodicity, directionality, and randomness: Wold features for image modeling and retrieval. *IEEE trans. on pattern analysis and machine intelligence*, vol. 18, no. 7, juillet 1996, pp. 722–733.

- [82] Lorette A., Descombes X. et Zerubia J. – Texture analysis through a Markovian modelling and fuzzy classification: Application to urban areas extraction from satellite images. – soumis à international journal of computer vision, juillet 1998. RR-Inria-3424 mai 1998.
- [83] Lorette A., Descombes X. et Zerubia J. – Texture analysis through Markov random fields: Urban areas extraction. *In: ICIP.* – Kobe, octobre 1999.
- [84] Ma W. Y. et Manjunath B. S. – Texture feature and learning similarity. *In: CVPR*, pp. 425–430. – San Francisco, juin 1996.
- [85] Malik J. et Perona P. – Preattentive texture discrimination with early vision mechanisms. *Journal of optical society of America*, vol. 7, no. 2, mai 1990, pp. 923–932.
- [86] Mao J. M. et Jain A. K. – Texture classification and segmentation using multiresolution autoregressive models. *Pattern recognition*, vol. 25, no. 2, 1992, pp. 173–188.
- [87] Maragos P. – Pattern spectrum and multiscale shape representation. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 11, no. 7, juillet 1989, pp. 701–716.
- [88] Matheron G. – *Éléments pour une théorie des milieux poreux.* – Masson, 1967.
- [89] Matheron G. – *Random sets and integral geometry.* – New-York, Wiley, 1975.
- [90] Matsuba I. – Renormalization group approach to hierarchical image analysis. *In: ICASSP*, pp. 1044–1047. – Seattle, mai 1988.
- [91] Minkowski H. – Volumen and oberfläche. *Math. Ann.*, vol. 57, 1903, pp. 447–495.
- [92] Mood A., Graybill F. et Boes D. – *Introduction to the theory of statistics.* – 1974.
- [93] Nguyen H. H. et Cohen P. – Gibbs random fields, fuzzy clustering, and the unsupervised segmentation of textured images. *CVGIP*, vol. 55, no. 1, janvier 1993, pp. 1–19.
- [94] Olivier C., Jouzel F., Matouat A. El et Courtellemont P. – Un nouveau critère pour la sélection de l'ordre d'un modèle. *In: GRETSI*, pp. 451–454. – Grenoble, 1997.

- [95] Palubinskas G., Descombes X. et Kruggel F. – An unsupervised clustering method using the entropy minimization. *In: ICPR*, pp. 1816–1818. – Australia, août 1998.
- [96] Pauc G., Latry C., Pasquet A. et Breton E. – Pourquoi et comment simuler des images SPOT 5. *Bul. S.F.P.T.*, no. 141, janvier 1996, pp. 22–27.
- [97] Peleg S., Naor J., Hartley R. et Avnir D. – Multiple resolution texture analysis and classification. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 6, no. 4, 1984, pp. 518–523.
- [98] Pentland A. P. – Fractal-based description of natural scenes. *IEEE Trans. on pattern analysis and machine intelligence*, vol. 6, no. 6, novembre 1984, pp. 661–672.
- [99] Picard R. W., Elfadel I. M. et Pentland A. P. – Markov/ Gibbs texture modeling: Aura matrices and temperature effects. *In: CVPR*, pp. 371–377. – Maui, juin 1991.
- [100] Prêteux F. et Descombes X. – Synthèse et analyse de textures par coopération de processus multi-échelles. *In: RFIA*, pp. 1015–1026. – Lyon, novembre 1991.
- [101] Richard F., Falzon F., Zerubia J. et Giraudon G. – Segmentation of SPOT images using Markov random fields. *In: EUSIPCO*, pp. 2493–2496. – Rhodes, septembre 1998.
- [102] Rissanen J. – Modeling by shortest data description. *Automatica*, vol. 14, 1978, pp. 465–471.
- [103] Rissanen J. – The minimal description length principle. – Research Report RJ-8236, juillet 1991. IBM US research centers.
- [104] Ruspini E. H. – A new approach to clustering. *Information and control*, vol. 15, no. 1, juillet 1969, pp. 22–32.
- [105] Salembier P. et Oliveras A. – Practical extensions of connected operators. *In: International symposium on mathematical morphology*, pp. 97–110. – Atlanta, mai 1996.
- [106] Schmitt M. et Mattioli J. – *Morphologie Mathématique*. – Masson, 1994.

- [107] Schroeder M. – *Fractals, chaos, power laws : minutes from an infinite paradise.* – W. H. Freeman and Company, 1990.
- [108] Serendero M. A. – *Extraction d'informations symboliques en imagerie SPOT: réseaux de communication et agglomérations.* – Thèse de Doctorat, UNSA, décembre 1989.
- [109] Serra J. – Morphologie mathématique et genèse des concrétions carbonatées des minerais de fer de Lorraine. *Sedimentology*, 1968, pp. 183–208.
- [110] Serra J. – *Image Analysis and Mathematical Morphology.* – Academic Press, 1982, vol. 1.
- [111] Serra J. – *Image Analysis and Mathematical Morphology.* – Academic Press, 1988, vol. 2: theoretical advances.
- [112] Serra J. et Soille P. – *Mathematical Morphology and Its Applications to Image Processing.* – Kluwer academic publishers, 1994.
- [113] Shannon C. E. – A mathematical theory of communication. *Bell systems technical journal*, vol. 27, no. 379, 1948.
- [114] Sigelle M. et Ronfard R. – Modèles de Potts et relaxation d'images de labels par champs de Markov. *Traitement du signal*, vol. 6, no. 9, 1992, pp. 449–458.
- [115] Sivakumar K. et Goutsias J. – Discrete morphological size distributions and densities : Estimation techniques and applications. *Journal of electronic imaging*, vol. 6, no. 1, janvier 1997, pp. 31–53.
- [116] Teuner A., Pichler O. et Hosticka B. J. – Unsupervised texture segmentation of images using tuned matched Gabor filters. *IEEE Trans. on image processing*, vol. 4, no. 6, juin 1995, pp. 863–870.
- [117] Tomita F. et Tsuji S. – *Computer Analysis of Visual Textures*, chap. 5- Structural Texture Analysis, pp. 71–82. – Kluwer Academic, 1990.
- [118] Tuceryan M. et Jain A. K. – *Handbook of Pattern Recognition and Computer Vision*, chap. 2.1-Texture Analysis, pp. 235–276. – World Scientific Company, 1993.
- [119] Unser M. – Texture classification and segmentation using wavelet frames. *IEEE Trans. on image processing*, vol. 4, no. 11, novembre 1995, pp. 1549–1560.

- [120] Van Laarhoven P. et Aarts E. – *Simulated annealing: Theory and applications*. – Reidel Pub., 1987.
- [121] Vincent L. – Morphological grayscale reconstruction: Definition, efficient algorithm and applications in image analysis. *In: CVPR*, pp. 633–635. – Urbana Champaign, juin 1992.
- [122] Vincent L. – Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Trans. on image processing*, vol. 2, no. 2, avril 1993, pp. 176–201.
- [123] Voiron C. – *Analyse spatiale et analyse d'image par morphologie mathématique*. – GIP Reclus publication, 1995.
- [124] Weber C. – *Images satellitaires et milieu urbain*. – 1995. Hermès Ed., Collection géomatique.
- [125] Weigl K. et Yu S. – A hybrid approach for image classification using neural network and Markov random field. *In: ICANN*, pp. 321–326. – Paris, octobre 1995.
- [126] Weldon T. P. et Higgins W. E. – Integrated approach to texture segmentation using multiple Gabor filters. *In: ICIP*, pp. 955–958. – Lausanne, septembre 1996.
- [127] Wilson K. et Kogut J. – The renormalization group and the  $\epsilon$  expansion. *Physics Report*, vol. 12, 1974, pp. 75–200.
- [128] Winkler G. – *Image analysis, random fields and dynamic Monte Carlo methods: a mathematical introduction*. – Springer-Verlag, 1995.
- [129] Winter A., Maître H., Cambou N. et Legrand E. – An original multi-sensor approach to scale-based image analysis for aerial and satellite images. *In: ICIP*, pp. 234–237. – Santa-Barbara, 1997.
- [130] Woods J. – Markov image modeling. *IEEE Trans. on automatic control*, vol. 23, no. 5, octobre 1978, pp. 846–850.
- [131] Younes L. – Estimation and annealing for Gibbsian fields. *Annales Institut Henri Poincaré*, vol. 24, no. 2, 1988, pp. 269–294.
- [132] Zadeh L. A. – Fuzzy sets. *Information and control*, vol. 8, 1965, pp. 338–353.

- [133] Zhu S.C., Wu Y. et Mumford D. – FRAME: filters, random fields, and minimax entropy towards a unified theory for texture modeling. *In: CVPR*, pp. 686–693. – San Francisco, juin 1996.
- [134] Zucker S. W. – Toward a model of texture. *CGIP*, no. 5, 1976, pp. 190–202.



## RÉSUMÉ

Dans cette thèse, nous nous intéressons au problème de l'analyse urbaine à partir d'images satellitales par des méthodes automatiques ou semi-automatiques issues du traitement d'image. Dans le premier chapitre, nous présentons le contexte dans lequel le travail a été effectué. Nous exposons les types de données utilisées, les approches statistiques considérées. Nous donnons également quelques exemples d'applications qui justifient une telle étude. Enfin, un état de l'art des diverses méthodes d'analyse des textures est présenté.

Dans les deux chapitres suivants, nous développons une méthode automatique d'extraction d'un masque urbain à partir d'une analyse de la texture de l'image.

Des méthodes d'extraction d'un masque urbain sont décrites. Ensuite, nous définissons plus précisément les huit modèles markoviens gaussiens fondés sur des chaînes. Ces modèles sont renormalisés par une méthode de renormalisation de groupe issue de la physique statistique afin de corriger le biais introduit par l'anisotropie du réseau de pixels. L'analyse de texture proposée est comparée avec deux méthodes classiques: les matrices de cooccurrence et les filtres de Gabor.

L'image du paramètre de texture est ensuite classifiée avec un algorithme non supervisé de classification floue fondée sur la définition d'un critère entropique. Les paramètres estimés avec cet algorithme sont intégrés dans un modèle markovien de segmentation. Des résultats d'extraction de masques urbains sont finalement présentés sur des images satellitales optiques SPOT3, des simulations SPOT5, et des images radar ERS1.

Dans le quatrième chapitre, nous présentons l'analyse granulométrique utilisée pour analyser le paysage urbain. Les outils et définitions de base de la morphologie mathématique sont exposés. Nous nous intéressons plus particulièrement à l'ouverture par reconstruction qui est utilisée comme transformation de base de la granulométrie. L'étape de quantification qui suit tout étape de transformation nous permet d'estimer en chaque pixel une distribution locale de taille qui est intégrée dans le terme d'attache aux données d'un modèle markovien de segmentation. Des tests sont effectués sur des simulations SPOT5.

**Mots clefs:** Analyse de texture, Segmentation, Classification floue, Renormalisation par décimation, Champs de Markov, Morphologie mathématique, Ouverture par reconstruction, Granulométrie, Distributions locales de taille, Observation de la terre, Analyse urbaine.

---

## ABSTRACT

In this thesis, we investigate the problem of urban areas analysis from satellite images by automatic or semi-automatic methods coming from image processing. In the first chapter, we describe the context of this work, i.e. the type of used data, the statistical applied methods. We also give some examples of the applications which require such an analysis. Finally, a study of the existing methods of texture analysis is presented.

In the second and third chapter, we develop a non supervised method based on texture analysis in order to extract an urban mask. First a study of the existing methods of urban mask extraction is presented. Second we precisely describe the eight chain-based Gaussian Markovian models used to characterize urban texture. These models are normalized through a renormalization group technique derived from statistical physics in order to correct the bias introduced by the anisotropy of the lattice. The above mentioned method of texture analysis is then compared with two classical ones: cooccurrences matrix and Gabor filters.

The image is then partitionned by an unsupervised fuzzy Cmeans algorithm based on an entropic criterion. The final segmentation is performed by the minimization of an energy derived from a Markovian model. Some results are presented that are obtained from SPOT3 images, SPOT5 simulations and radar ERS1 images.

In the fourth chapter, we present the granulometric approach used to segment within the urban area itself. The basic operations and definitions of mathematical morphology are settled. We are particularly interested in opening by reconstruction operation based on geodesic dilatations. In fact this operation is used to define a granulometry. The quantification step that follows the transformation step consists in estimating a local size distribution function for each pixel. These parameters are then integrated in the data term of a Markovian model. Some results on SPOT5 simulations are presented.

**Keywords:** Texture analysis, Segmentation, Fuzzy classification, Renormalization by decimation, Markov random fields, Mathematical morphology, Opening by reconstruction, Granulometry, Local size distribution densities, Earth observation, Urban areas analysis.