



HAL
open science

Méthode collaborative de segmentation et classification d'objets à partir d'images de télédétection à très haute résolution spatiale

Aymen Sellaouti

► **To cite this version:**

Aymen Sellaouti. Méthode collaborative de segmentation et classification d'objets à partir d'images de télédétection à très haute résolution spatiale. Traitement du signal et de l'image [eess.SP]. Université de Strasbourg; Université Tunis El Manar. Faculté des Sciences Mathématiques, Physiques et Naturelles de Tunis (Tunisie), 2014. Français. NNT : 2014STRAD032 . tel-01271036

HAL Id: tel-01271036

<https://theses.hal.science/tel-01271036>

Submitted on 8 Feb 2016

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

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



UNIVERSITÉ DE TUNIS EL MANAR

FACULTÉ DES SCIENCES DE TUNIS



Université de Strasbourg

**Laboratoire des sciences de l'ingénieur, de
l'informatique et de l'imagerie**

THÈSE

présentée en vue de l'obtention du

Diplôme de Docteur en Informatique

par

Aymen SELLAOUTI

**Méthode collaborative de segmentation et
classification d'objets à partir d'images de
télédétection à très haute résolution
spatiale.**

Soutenue publiquement le 16 Septembre 2014

Membres du jury

Directeur de thèse - M. Khaled BSAÏES, Professeur, FST, Université Tunis El Manar

Directeur de thèse - Mme Aline DERUYVER, Maître de conférences Université de Strasbourg (France)

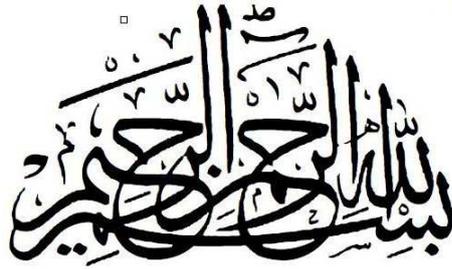
Président de Jury - M. Yahia SLIMANI, Professeur, ISAM, Université de la Manouba

Rapporteur - M. Nicolas LOMENIE, Maître de conférences, Université Paris Descartes (France)

Rapporteur - M. Imed Riadh FARAH, Maître de conférences, Université de la Manouba

Examineur - M. Pierre COLLET, Professeur, Université de Strasbourg (France)

Préparée sous convention de cotutelle UTM-FST (Tunisie) – UDS-Strasbourg (France)



Dédicaces

Je voudrais tout d'abord remercier infiniment Dieu.

En témoignage de ma gratitude, de mon amour et de ma grande reconnaissance, je dédie ce travail à :

L'âme de mes grands-parents Habiba Chedlia et Ahmed, qu'ils reposent en paix, vous serez toujours dans mon coeur.

Mon grand-père Romdhan le doyen de notre famille et son guide

Ma mère Nebiha qui a toujours été là pour moi et je la remercie pour son amour, sa patience et sa confiance en moi.

Mon père Abderrahmen pour son soutien illimité et pour la confiance qu'il a mise en moi.

Ma futur épouse Rym qui m'a soutenu et poussé dans la dernière ligne droite.

Ma soeur Monia et son époux Yassin,

Toute la famille Sellaouti,

Toute la famille Ben Tamarzist,

Mes amis,

Mes enseignants à la Faculté des Sciences de Tunis.

Toutes les personnes qui ont fait de moi ce que je suis maintenant.

Tous les gens que j'aime.

Remerciements

Je tiens tout d'abord à remercier Mr. Yahya SLIMANI, Professeur à l'Université de la Manouba, pour l'honneur qu'il me fait en présidant ce jury.

Je remercie également Messieurs Pierre COLLET, Professeur à l'Université de Strasbourg, Nicolas LOMENIE, Maître de Conférences HDR à l'Université Paris Descartes et Imed Riadh FARAH, Maître de Conférences HDR à L'Université de la Manouba d'avoir bien voulu participer à ce jury et juger mon travail.

Je tiens à adresser mes vifs remerciements et ma profonde reconnaissance à Madame Aline DERUYVER, Maître de Conférences HDR à l'Université de Strasbourg, et Mr Khaled BSAIES, Professeur à la Faculté des Sciences de Tunis, pour avoir dirigé mon travail, pour leurs assistance, conseils et disponibilité, pour leurs encouragements et leurs soutiens continus et pour l'intérêt qu'ils ont manifesté pour mon travail, leurs suggestions, encouragements et remarques ont été d'une importance capitale.

Je remercie aussi Mr Atef HAMOUDA, Maître Assistant à la Faculté des Sciences de Tunis, l'homme qui a motivé dès le premier jour d'enseignement à l'Amphi C mon envie d'apprendre et de devenir un jour une personne et un enseignant comme lui. Merci " mon deuxième père " d'avoir été toujours là pour moi et de m'avoir toujours poussé à avancer et à croire en moi et d'avoir toujours gardé votre porte ouverte pour moi.

Je remercie aussi Mr Cedric WEMMERT, Maître de Conférences HDR à l'université de Strasbourg, sans qui cette thèse ne serait pas finalisée. Merci pour votre aide et vos précieuses remarques et suggestions et pour le savoir que vous m'avez transmis. Ce fût un plaisir énorme de vous côtoyer à chaque passage à Strasbourg et d'avoir travaillé à vos côtés.

Merci à Mr Sadok BEN YAHIA, Professeur à la Faculté des Sciences de Tunis, pour

son aide si précieuse, ses consignes et son savoir-faire qu'il a toujours partagé avec moi.

Un grand merci pour mon ami et frère Hmida ROJBANI, pour les longues discussions qu'on a eu et qui ont été une source d'inspiration énorme pour l'avancement de cette thèse. Merci pour l'aide continue que tu m'as apporté et pour avoir été toujours là pour moi.

Merci à Nidhal JELASSI pour toutes ces années passées ensemble à suivre ce même chemin et ce même but que nous atteignons encore une fois en même temps et ensemble. Le destin a fait que nos chemins se croisent dès la première année de faculté et nous voilà des années après encore ensemble pour terminer ce voyage vers la thèse.

Merci à Mohamed NAOUAI et Zouhour BEN DHIAF pour leurs soutiens.

Merci à tous les membres du laboratoire LIPAH pour ces années passées ensemble et pour tout ce que vous avez apporté pour la finalisation de ce travail.

Table des matières

Remerciements	iii
1 L'analyse d'images satellitaires	8
1.1 Introduction	8
1.2 Analyse d'images basée pixels	9
1.2.1 Principe	9
1.2.2 Classification	12
1.2.2.1 Définition	12
1.2.2.2 Classification supervisée	12
1.2.2.3 Classification non supervisée	13
1.3 Du pixel vers l'objet	13
1.4 Analyse d'images basée objets	18
1.4.1 Principe des approches basées objets	18
1.4.2 Segmentation d'images	20
1.4.2.1 Segmentation par pixel (par seuillage)	21
1.4.2.2 Segmentation par contour	21
1.4.2.3 Segmentation par région	21
1.4.2.4 Segmentation par collaboration région/contour	22
1.4.3 Caractérisation des objets	22
1.4.4 Identification des objets	23
1.4.4.1 Les méthodes par attributs	24

1.4.4.2	Les méthodes basées connaissances	24
1.5	Etat de l'art et positionnement	25
1.5.1	Approche naturelle	25
1.5.2	Approche hiérarchique	28
1.5.3	Approche hybride	30
1.5.4	Limite des approches d'analyse d'images basée objets	33
1.6	Conclusion	34
2	Croissance de région hiérarchique sémantique : une approche collaborative pour la segmentation et la classification d'objets.	35
2.1	Introduction	35
2.2	Croissance de régions : État de l'art et positionnement	36
2.2.1	Croissance de régions bas niveau	37
2.2.2	Croissance de régions sémantique	40
2.2.3	Synthèse et étude comparative	43
2.3	Principe de l'approche proposée	45
2.3.1	Prétraitement	47
2.3.1.1	Segmentation	47
2.3.1.2	Classification	47
2.3.2	Croissance hiérarchique sémantique	48
2.3.2.1	Extraction des germes	50
2.3.2.2	Intégration des connaissances spatiales d'adjacence	51
2.3.2.3	Croissance sémantique	53
2.4	Conclusion	57
3	Croissance hiérarchique sémantique pour l'extraction des bâtiments et des routes	58
3.1	Introduction	58
3.2	Connaissances du domaine	58

3.2.1	Le fossé sémantique	58
3.2.2	Dictionnaire des données	59
3.2.3	Formalisme des données	61
3.2.3.1	Hierarchie de classes	61
3.2.3.2	Caractérisation des classes	62
3.3	Template dynamique pour l'extraction des bâtiments	63
3.3.1	Méthodes objets pour l'extraction des bâtiments	63
3.3.2	Template dynamique pour l'extraction de la zone de croissance .	64
3.3.3	Croissance sémantique pour l'extraction des bâtiments	69
3.3.3.1	Aire	69
3.3.3.2	Rectangularité	71
3.3.3.3	Homogénéité	74
3.4	Extraction des routes basée sur la transformé de Radon	75
3.4.1	Méthodes objets pour l'extraction des routes	75
3.4.2	Transformée de Radon	77
3.4.2.1	Transformée de Radon continue	77
3.4.2.2	Transformée de Radon discrète	77
3.4.3	La transformée de Radon pour l'estimation de la direction des routes	78
3.4.4	Extraction des routes	81
3.5	Expérimentation et évaluation	84
3.6	Conclusion	86
4	Système multi-agents évolutionnaire pour une approche coopérative régions-contours pour l'extraction des objets d'intérêt	88
4.1	Introduction	88
4.2	Les approches de segmentation coopératives régions-contours : état de l'art	89
4.3	Les systèmes multi-agents	91

4.3.1	Les agents	91
4.3.1.1	Définition	91
4.3.1.2	Types d'agents	92
4.3.2	Interaction au sein des systèmes multi-agents	93
4.3.3	Les systèmes multi-agents dans l'analyse d'images	94
4.4	Système multi-agents collaboratif régions-contours	96
4.4.1	Architecture du système multi-agents	97
4.4.2	Prétraitement	97
4.4.3	Extraction des objets.	97
4.4.4	Les agents composant le système multi-agents	97
4.4.4.1	L'agent superviseur	99
4.4.4.2	L'agent contour	99
4.4.4.3	L'agent région	100
4.4.5	Critique du système multi-agents	100
4.5	Système multi-agents auto-adaptatif basé sur les algorithmes génétiques	101
4.5.1	Les algorithmes génétiques	101
4.5.1.1	Génération de la population initiale	102
4.5.1.2	Principes de sélection	102
4.5.1.3	Opérateur de croisement	103
4.5.1.4	Opérateur de mutation	105
4.5.1.5	Fonction de fitness	105
4.5.1.6	Condition d'arrêt	105
4.5.2	Architecture du système multi-agents auto-adaptatif	106
4.5.3	Apport de l'application de l'algorithme génétique dans l'amélioration de notre système multi-agents	107
4.6	Conclusion	107
5	Un système multi-agents évolutionnaire auto-adaptatif pour l'extraction des bâtiments	109

5.1	Introduction	109
5.2	Un système multi-agents pour l'extraction des bâtiments	109
5.2.1	Prétraitement	111
5.2.2	Extraction des bâtiments	113
5.3	Algorithme évolutionnaire local automatisant le SMA pour l'extraction des bâtiments	116
5.3.1	Codage des individus	116
5.3.2	Fonction de fitness	116
5.3.3	Opérateurs génétiques utilisés	118
5.3.3.1	L'opérateur de sélection	118
5.3.3.2	L'opérateur de croisement	119
5.3.3.3	L'opérateur de mutation	119
5.3.4	Fonctionnement de l'algorithme génétique	119
5.3.4.1	Création de la population initiale	120
5.3.4.2	Évaluation de la population	120
5.3.4.3	Évolution des populations	120
5.4	Expérimentation et évaluation	120
5.5	Conclusion	125
6	Validation et expérimentation dans le cadre de l'analyse de scène urbaine	127
6.1	Introduction	127
6.2	Données utilisées	127
6.3	Métriques d'évaluation	134
6.4	Évaluation de la classe Bâtiment	135
6.5	Évaluation de la classe Route	141
6.6	Conclusion	147
	Conclusion générale et perspectives	148

Liste des figures

1.1	Principe de la classification basée pixels.	9
1.2	La classification basée pixels : une classification statistique basée sur les données spectrales.	10
1.3	Exemple d'analyse basée pixels sur une image artificielle.	11
1.4	Illustration de l'apport de l'information contextuelle dans l'analyse des images [Lang <i>et al.</i> , 2009]	14
1.5	Workflow de la comparaison des approches objet et pixel.	17
1.6	Exemple d'analyse d'image basée objets sur une image artificielle	19
1.7	Graphe de référence du cerveau humain.	26
1.8	Workflow de l'approche naturelle proposée par Walter [Walter, 2004]. . .	27
1.9	Vision hiérarchique de l'image : chaque objet image est lié à son super objet et son sous objet.	29
1.10	Workflow de l'approche hiérarchique [Aksoy et Ercanoglu, 2012].	31
2.1	Architecture de l'approche sémantique de Fan <i>et al.</i> [Fan <i>et al.</i> , 2001b].	41
2.2	Architecture de la croissance sémantique d'Athanasiadis <i>et al.</i> [Athanasiadis et Kollias, 2007].	42
2.3	Workflow de l'approche.	46
2.4	Diagramme de croissance de régions sémantique.	50
2.5	Illustration de l'extraction des germes sur une image artificielle.	52

2.6	Illustration de l'intégration des contraintes dans la croissance hiérarchique sur une image artificielle.	56
3.1	Extrait du dictionnaire des données présentant une description visuelle de la classe "bâtiment : habitat collectif".	61
3.2	Extrait de la hiérarchie de classes utilisée dans [Forestier, 2010].	62
3.3	Description visuelle des pavillons.	65
3.4	Création du template.	66
3.5	Illustration des templates élagués.	67
3.6	Création du template dynamique.	68
3.7	Croissance par meilleur ajustement.	70
3.8	Allure de la courbe de la fonction <i>Aire</i>	71
3.9	Différence entre la <i>GR – signature</i> d'un objet rectangulaire parfait et la <i>GR – signature</i> d'un objet rectangulaire non parfait.	74
3.10	Projection selon des rayons parallèles à l'angle θ_0 [Wang et Li, 2006].	78
3.11	Plongement du domaine f dans une grille discrète.	79
3.12	Illustration graphique de la classe "Route", extrait du dictionnaire des données.	79
3.13	Extraction de la fenêtre centrée sur le germe pour l'application de la transformée de Radon.	80
3.14	Résultat de l'application de la transformée de Radon.	81
3.15	Résultat de la projection sur trois angles différents.	81
3.16	La direction détectée par l'application de la transformée de Radon.	82
3.17	Extraction de la zone de croissance pour un germe de type route.	83
3.18	Routes linéaires par morceaux.	83
3.19	Zones urbaines de la ville de Strasbourg	85
3.20	Extraction de la classe "Bâtiment"	86
3.21	Extraction de la classe "Route"	86
4.1	Segmentation contours-régions séquentielle.	90

4.2	Segmentation contours-régions par coopération de résultats.	90
4.3	Segmentation mutuelle contours-régions.	91
4.4	Fonctionnement des agents réactif.	92
4.5	Fonctionnement des agents cognitif.	93
4.6	Interaction inter-agents.	94
4.7	Architecture du système multi-agents proposé.	98
4.8	Coopération séquentielle.	99
4.9	Algorithme génétique.	101
4.10	Schéma d'une roulette	103
4.11	Croisement en 1 point	104
4.12	Croisement en 2 points	104
4.13	Principe de l'opérateur de mutation	105
4.14	Schéma global de notre système multi-agents intégrant l'algorithme génétique.	106
5.1	Architecture du SMA pour l'extraction des bâtiments.	110
5.2	Phase de prétraitement pour l'extraction des bâtiments.	112
5.3	Coopération entre un agent région et un agent contour pour l'extraction d'un bâtiment.	113
5.4	La distance de Bhattacharya.	114
5.5	SMA Évolutionnaire local pour l'extraction des bâtiments.	117
5.6	Les images de test.	121
5.7	Images proche infrarouge des deux sites.	122
5.8	Résultat de l'indice de végétation des images test	122
5.9	Résultat de l'opérateur de dilatation sur les images test	123
5.10	Résultat de l'application de l'opérateur de Harris sur les images test.	123
5.11	Extraction de la classe "Bâtiment"	124
5.12	Extraction des bâtiments au cours de l'évolution de l'algorithme génétique.	125
5.13	Évolution du rappel et de la précision.	125

6.1	Vue spatiale du satellite Quickbird	128
6.2	Illustration de l'image intégrale de la zone d'étude en couleur	129
6.3	Illustration de l'image intégrale de la zone d'étude en proche infra rouge (PIR)	130
6.4	Image de test de Cronenbourg ouest.	131
6.5	Image de test de La Meinau.	132
6.6	Images de test d'Oberhausbergen.	132
6.7	Images de test du Bas Rhin.	133
6.8	Images de test d'Eckbolsheim.	133
6.9	Résultat de l'extraction des bâtiments dans les images de test I.	136
6.10	Résultat de l'extraction des bâtiments dans les images de test II.	137
6.11	Résultat de l'extraction des bâtiments dans les images de test III.	138
6.12	Comparaison entre les performances des deux approches.	140
6.13	Comparaison entre les performances de différentes approches d'extraction de bâtiments.	141
6.14	Résultat de l'extraction des routes dans les images de test (...)	142
6.15	Résultat de l'extraction des routes dans les images de test (suite).	143
6.16	Image représentant un rond-point dans un site très dense.	144
6.17	Image représentant un rond-point dans un site éparse.	145
6.18	Résultats d'extraction des ronds-point.	145
6.19	Comparaison entre les performances de différentes approches d'extraction de routes.	146

Liste des tableaux

2.1	Étude comparative des méthodes de croissance de régions	44
3.1	Description des bâtiments (Dictionnaire de données Fodomust)	64
3.2	Évaluation du rappel et de la précision	85
5.1	Évaluation du rappel et de la précision du SMA évolutionnaire pour la classe "Bâtiment"	124
5.2	Evolution de l'algorithme génétique	125
6.1	Informations sur le satellite Quickbird	128
6.2	Caractéristiques des images de test	134
6.3	Évaluation quantitative de la classe bâtiment.	139
6.4	Performance des approches d'extraction de bâtiments	140
6.5	Evaluation quantitative de la classe route	141
6.6	Performance des approches d'extraction de routes	146

Introduction générale

Mise en contexte

L'interprétation des images satellitaires devient de plus en plus complexe avec l'avènement des images satellitaires à Très Haute Résolution Spatiale (THRS). En effet, l'augmentation rapide de la résolution spatiale (taille de l'image) ainsi que de la résolution spectrale (nombre de bandes) accroît la complexité des images disponibles. L'évolution récente des technologies d'observation de la Terre permet une acquisition massive de nouvelles images et rend ainsi disponible des volumes de données importants. Ces images présentent un potentiel important pour l'acquisition de l'information géographique et pour la cartographie du territoire.

Les méthodes d'analyse conventionnelles dites pixéliques ont montré leurs limites face à cette évolution. La croissance des détails dans l'image ainsi que la richesse des informations qui s'y trouvent font que les approches pixéliques trouvent énormément de difficultés à produire des résultats convenables et qui peuvent satisfaire les besoins des utilisateurs. Afin de palier à ces difficultés, plusieurs recherches ont été menées pour se familiariser avec les nouvelles données de la très haute résolution, exploiter la richesse des informations présentes dans l'image et intégrer des connaissances experts. C'est dans ce contexte qu'est né l'analyse d'image basée objets [Blaschke, 2010]. Cette méthode d'analyse objet a permis de passer d'une vision pixélique de l'image vers une vision objet, ainsi, la plus petite entité de l'image n'est plus le pixel mais un objet (i.e., région) composé d'un ensemble de pixels homogènes selon un critère donné. Cette modélisation de l'image en un ensemble d'objets permet une meilleure description et une meilleure interprétation de l'image. En effet, les objets présentent d'avantages de caractéristiques spectrales que les pixels (la valeur moyenne par bande, la valeur minimale, maximale et médiane, la variance, ...). Elles présentent aussi des paramètres

géométrique (le périmètre, l'air, la longueur, la largeur, ...) et spatiale.

De nombreux travaux s'intéressent actuellement à la définition de méthodes basées objets pour l'extraction d'objets particuliers dans les images de télédétection à très haute résolution spatiale [Blaschke et Strobel, 2001, Blaschke et Lang, 2006, Derivaux *et al.*, 2010, Forestier *et al.*, 2008b, Blaschke, 2010]. Tous s'accordent à dire que cette extraction ne peut être faite de manière totalement non supervisée, c'est-à-dire, sans utilisation de connaissances du domaine sur les objets à extraire, ces connaissances pouvant prendre plusieurs formes : ontologies du domaine, contraintes, relations inter-objets, etc.

Des travaux dans cet axe de recherche ont fait l'objet d'une collaboration entre le laboratoire d'Informatique Icube et le laboratoire de Géographie Environnementale (LIVE) de l'université de Strasbourg [Derivaux, 2009, Forestier, 2010, Kurtz, 2012].

Dans ces travaux, Dérivaux [Derivaux, 2009] s'est penché sur deux axes de recherche. Le premier axe vise à intégrer les connaissances du domaine dans les approches d'analyse d'images orientée objet. Il a présenté plusieurs méthodes d'intégration des connaissances expertes dans les deux étapes de l'analyse d'images orientée objet, à savoir, l'étape d'identification et l'étape de construction. Il a proposé plusieurs méthodes d'intégration des connaissances, qu'il s'agisse d'exemples ou de connaissances du domaine. Dans le deuxième axe, Il a aussi entamé des travaux sur la mise en cause de l'étape d'identification par les résultats de l'étape de construction en utilisant un algorithme génétique permettant le paramétrage de la segmentation. Et c'est dans cet axe là que se situe nos travaux qui présente une continuité des travaux de Dérivaux.

Forestier [Forestier, 2010] s'est concentré sur les algorithmes de clustering et plus particulièrement les algorithmes de clustering collaboratifs en étudiant plus particulièrement la combinaison des résultats de clustering ainsi que la combinaison des algorithmes de clustering. Il s'est aussi intéressé à l'étude de l'intégration des connaissances dans les algorithmes de clustering collaboratifs.

Finalement, Camille Kurtz [Kurtz, 2012] s'est intéressé à deux axes de recherches. Le premier axe concerne le développement d'approches objets non supervisées permettant l'extraction des objets d'intérêts à partir d'images multirésolutions. Le deuxième axe a visé l'intégration des connaissances expertes dans les processus d'analyse proposés dans le premier axe.

Problématique de la recherche

La multiplication des capteurs à Très Haute Résolution (spatiale ou spectrale) est une véritable opportunité pour l'identification des objets, notamment en milieu urbain. Toutefois, l'hétérogénéité de ces données satellites de résolution métrique (surfaces hétérogènes, ombres portées, etc) perturbe les méthodes de classification classique, dite spectrale, adaptées à l'analyse d'image de résolution décimétrique. En effet, et jusqu'à la fin des années 90, les approches d'analyse d'images de télédétection étaient fondées sur l'étude de l'image comme étant un ensemble de pixels ayant des propriétés bas niveau et plus généralement radiométrique permettant de distinguer l'ensemble de zones d'intérêts de l'image. La résolution spatiale de l'époque variant de la basse résolution (i.e. supérieur à $100M$) à la moyenne résolution (i.e. variant entre $30M$ et $100M$) permettait d'identifier des zones d'intérêts assez grossières telles que les parcelles agricoles.

Cependant, avec l'apparition des images à haute (i.e. variant entre $10M$ et $30M$) et très haute résolution (i.e. inférieur à $10M$), les travaux destinés aux approches basses et moyennes résolution ne sont plus adaptés. En effet, quoi qu'elles permettent d'offrir dans certains cas un résultat qui reflète de manière fidèle la réalité du terrain étudié, ces approches restent tributaires d'une conformité et d'une répartition homogène des classes recensées, ce qui n'est plus le cas avec les images actuelles. Plusieurs travaux ont montré les limites de ces approches [Gao et Mas, 2008, Robert *et al.*, 2010, Myint *et al.*, 2011]. La tendance actuelle est au développement de méthodes basées "objets" où chaque objet représente un ensemble de pixels homogènes et est construit sur les caractéristiques élémentaires (spectrale, géométrique) et les relations spatiales entre les objets (texture et relations contextuelles).

Plusieurs travaux ont été entamés sur l'utilisation de connaissances du domaine dans un processus de classification d'image. Cependant, dans les travaux actuels, les deux étapes de construction des objets (segmentation) puis de classification sont toujours faites de manière indépendante [Karsenty *et al.*, 2007, Jacquin *et al.*, 2008, Aminipouri *et al.*, 2009, Aksoy et Ercanoglu, 2012, Mahmoudi *et al.*, 2013]. De fait, si un objet a été mal construit/extrait lors de l'étape de segmentation, il ne pourra pas être classé correctement, car il ne correspondra pas (spatialement et spectralement) aux caractéristiques attendues pour ce type d'objet. Le manque d'interaction entre ces deux étapes fait que les résultats obtenus par la segmentation sont la pierre angu-

laire de tous le processus vu qu'elles ne sont pas remises en cause. Or, en étudiant de près la segmentation, nous remarquons qu'elle souffre d'un problème majeur, à savoir, qu'elle est un problème complexe mal posé. Ceci est du au fait que plusieurs solutions peuvent être générées et qu'il est difficile de dire laquelle des solutions trouvées est la meilleure. Cet inconvénient combiné avec le manque d'interaction entre cette étape de construction et l'étape d'identification des objets présentent un gros handicap pour les approches d'analyse d'image orientée objets.

Afin de palier à cet inconvénient, des travaux ont été entamés au sein de notre équipe. Derivaux a proposé une approche qui remet en cause les résultats de la segmentation initiale en se basant sur une évaluation du résultat de la classification obtenue [Derivaux, 2009]. Cette approche consiste à utiliser un algorithme évolutionnaire dans le but d'optimiser les paramètres de l'algorithme de segmentation initial. A chaque itération, les auteurs évaluent la classification obtenue à partir de la segmentation relative au jeu de paramètres introduit. Les paramètres finaux seront ceux qui auront maximisé la fonction d'évaluation. Cependant, cette approche permet certes d'avoir un bon paramétrage, mais ce paramétrage reste fondé sur une base de test et dépendant des propriétés des images sur lesquelles l'évaluation de la classification a été faite. Cette thèse représente la continuité de ces travaux.

Objectif

L'objectif principal de cette thèse est donc de remédier au problème de dépendance des approches d'analyse d'images orientées objets à la phase initiale de construction des objets. Les objectifs spécifiques autour desquelles s'est fondé ce travail sont la conception et le développement d'approches collaboratives d'analyse d'images orientées objets permettant :

- une interaction et une coopération entre l'étape de construction et d'identification des objets d'intérêts de l'image,
- l'intégration des connaissances du domaine et des contraintes qui peuvent être introduites par les experts.

Structure du document

Ce document est divisé en trois parties. La première partie introduit le contexte de notre travail (chapitre 1). Les deux autres parties représentent les deux approches que nous avons introduites (chapitre 2 et 3 & chapitre 4 et 5). La structure du présent rapport est la suivante :

Chapitre 1

Le premier chapitre introduit les approches de l'analyse d'images. Nous commençons par présenter les approches basées pixels et les limites qu'elles présentent par rapport à l'avènement des images à haute et très hautes résolutions. Nous présentons ensuite les approches d'analyse d'images basées objets comme solution à ces limites et l'intérêt croissant qu'elles suscitent dans la littérature. Nous étudions les différentes approches objets introduites dans la littérature et nous mettons l'accent sur l'inconvénient commun que présentent ces approches, à savoir, le manque d'interaction entre les deux étapes majeures du processus. En se fondant sur cet inconvénient, nous introduisons finalement les deux approches que nous détaillons dans les prochains chapitres.

Chapitre 2

Une réflexion sur les approches de croissance de région nous amène à introduire notre première approche. Cette approche est une approche d'analyse d'image orientée objets collaborative. Basé sur une architecture itérative, cette approche permet d'introduire la notion de la sémantique dans un processus de croissance de région. Une première étape de prétraitement permet de décomposer l'image en un ensemble d'objets homogènes. A chacun de ces objets, nous associons un score fondé sur les connaissances du domaine et les propriétés bas niveau de l'image. A partir de ces scores, nous itérons sur un processus de croissance en favorisant les objets qui maximisent ces scores. La croissance à partir d'un ensemble de régions sélectionnées pour chaque itération est une croissance sémantique spécifique à la classe d'appartenance de cette région. La hiérarchie permet aussi d'intégrer au fur et à mesure un ensemble de contraintes composé des régions déjà traitées.

Chapitre 3

Au cours de ce chapitre, nous introduisons deux approches de croissance sémantique que nous avons abordé dans le chapitre précédent. En effet, la phase de croissance sémantique reste tributaire de la classe d'appartenance de la région traitée. Nous nous sommes donc intéressés aux deux classes les plus visées pour les images de zones urbaines, à savoir, les routes et les bâtiments. En se basant sur les connaissances que nous possédons sur l'image à traiter. Nous avons proposé une première approche utilisant la transformée de Radon afin de détecter la direction de la zone de croissance de l'objet traitée. Cette information combinée avec les propriétés géométriques de la route a permis de limiter l'espace de croissance. Pour la classe bâtiment, les templates dynamiques créés à partir de connaissances géométriques ont permis de gérer l'espace de croissance. La croissance finale a été réalisée en utilisant la technique de meilleur ajustement en respectant un ensemble de contraintes spécifiques à la classe.

Chapitre 4

Ce chapitre aborde les approches de croissances de région collaboratives régions-contours comme base de l'analyse d'image orientée objets. Cette approche étant une collaboration entre les deux approches région et contour, nous introduisons un système multi-agents qui s'adapte à cette collaboration en y intégrant les avantages du multi-agents. L'intégration des connaissances lors de la phase de la construction permet de fusionner les deux étapes de construction et d'identification. Le paramétrage des deux algorithmes région et contour reste encore un inconvénient à surmonter. Afin d'y remédier, nous introduisons un algorithme évolutionnaire local permettant une automatisation et une spécification du paramétrage de ces deux algorithmes pour chaque objet d'intérêt traité.

Chapitre 5

Dans ce chapitre, nous proposons une adaptation de notre approche multi-agent évolutionnaire pour l'extraction des bâtiments. Nous introduisons l'ensemble des agents permettant l'acheminement des différentes étapes de notre système. Nous présentons aussi la modélisation de l'ensemble des connaissances expertes concernant les bâtiments dans la fonction de fitness permettant de gérer le paramétrage de chaque région pour

l'extraction finale du bâtiment.

L'analyse d'images satellites

1.1 Introduction

L'analyse des images satellites est en pleine évolution. En effet, avec l'avènement des nouvelles techniques d'acquisition d'images et le passage vers la très haute résolution, les approches d'analyse d'images classiques (basées pixels) ne donnent plus entière satisfaction. En effet, avec l'augmentation du bruit, des détails et la richesse d'informations de ces images, les approches basées pixels n'y sont plus adaptées.

Dans ce chapitre, nous introduisons, tout d'abord, les approches d'analyse d'images basées pixels ainsi que leurs limites pour l'analyse des images à très haute résolution. Nous présentons ensuite, le principe et les différentes étapes qui forment les approches d'analyse d'images basées objets. Ces approches se présentent comme une alternative permettant de pallier certaines limites des approches basées pixels. Plusieurs travaux ont été présentés dans la littérature. Nous proposons de décomposer ces approches en trois classes : approche naturelle, approche hiérarchique et approche collaborative. Nous présentons un aperçu de ces travaux. Finalement, nous illustrons les limites de ces approches sur lesquelles nous nous sommes focalisés pour introduire notre solution.

1.2 Analyse d'images basée pixels

1.2.1 Principe

Les approches basées pixels traitent l'image comme étant un ensemble de pixels. Le pixel est donc considéré comme la donnée élémentaire à traiter dans l'image. Plusieurs approches basées pixels ont été proposées en télédétection, ces approches se sont avérées très efficaces en ce qui concerne l'analyse des images basse et moyenne résolution. Comme l'illustre la figure 1.1, la majorité des approches basées pixels traite l'image comme un ensemble de matrices représentant chacune une bande spectrale. Les attributs décrivant les pixels sont extraits directement de l'image représentant des informations bas niveau et donc des grandeurs physiques telles que la radiométrie.

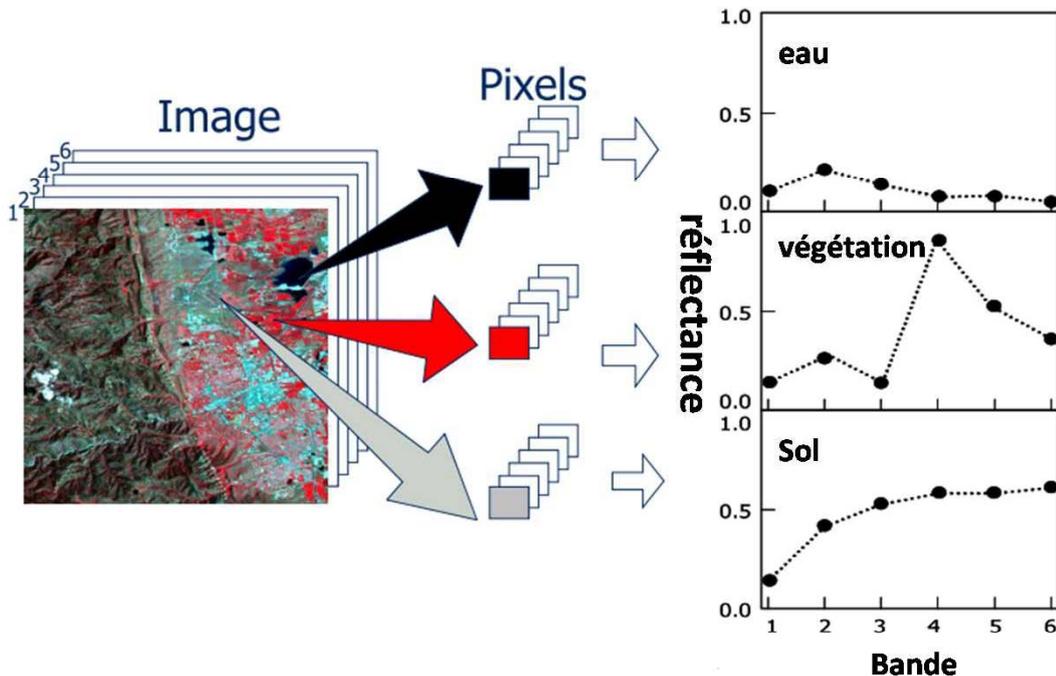


Figure 1.1 — Principe de la classification basée pixels.

Les pixels sont alors considérés comme un ensemble d'individus comme l'illustre la figure 1.2. Ces individus sont ensuite traités par des techniques statistiques conventionnelles comme les techniques de classification supervisée et non supervisée.

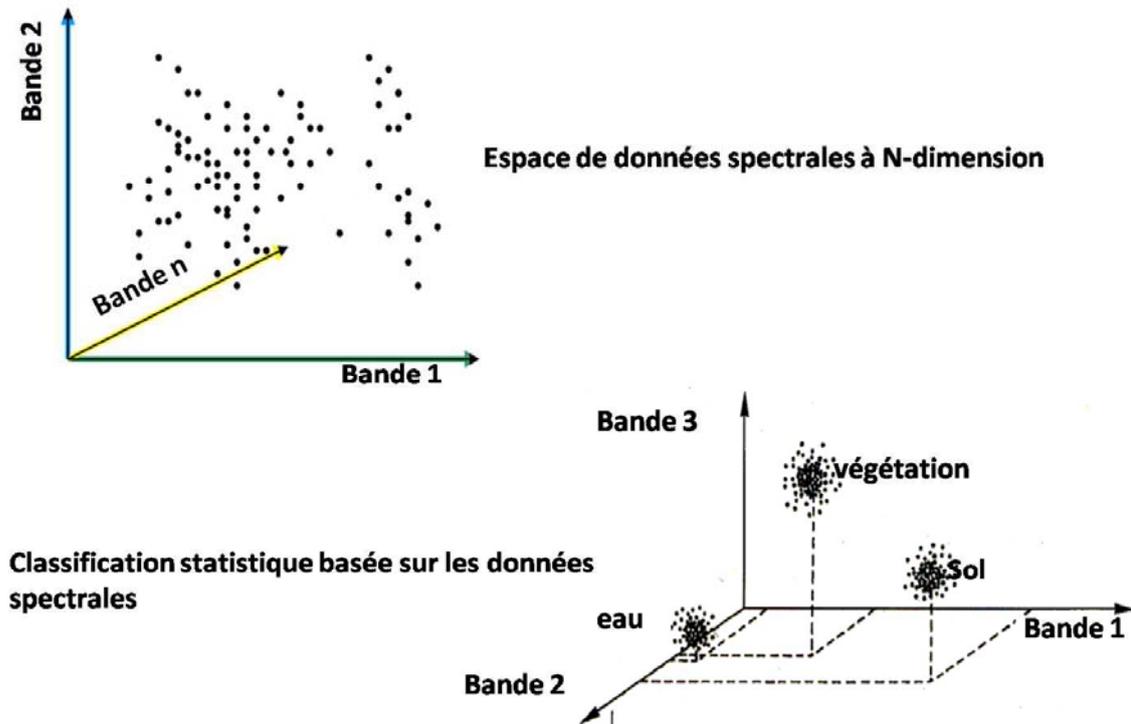


Figure 1.2 — La classification basée pixels : une classification statistique basée sur les données spectrales.

La figure 1.3 résume les étapes de l'analyse d'images basée pixels.

Plusieurs travaux ont utilisé les méthodes d'analyse d'images basées pixels afin de détecter des objets d'intérêt dans l'image. Gao et Wu proposent une classification non supervisée des quatre bandes spectrales d'une image Ikonos de la ville d'Auckland en Nouvelle Zélande comme la base d'un traitement intégrant le raisonnement spatial afin d'extraire le réseau routier [Gao et Wu, 2004].

Zhang et Wang utilisent une fusion d'images entre une image multi spectrale et une image panchromatique Quickbird. Ensuite, une classification non supervisée est appliquée afin d'extraire la classe route. Une détection de contour est appliquée sur l'image panchromatique permettant d'extraire la carte des contours qui permettra de segmenter l'image de la classe route. Les objets non routes sont ensuite supprimés par un algorithme de filtrage [Y. Zhang, 2004].

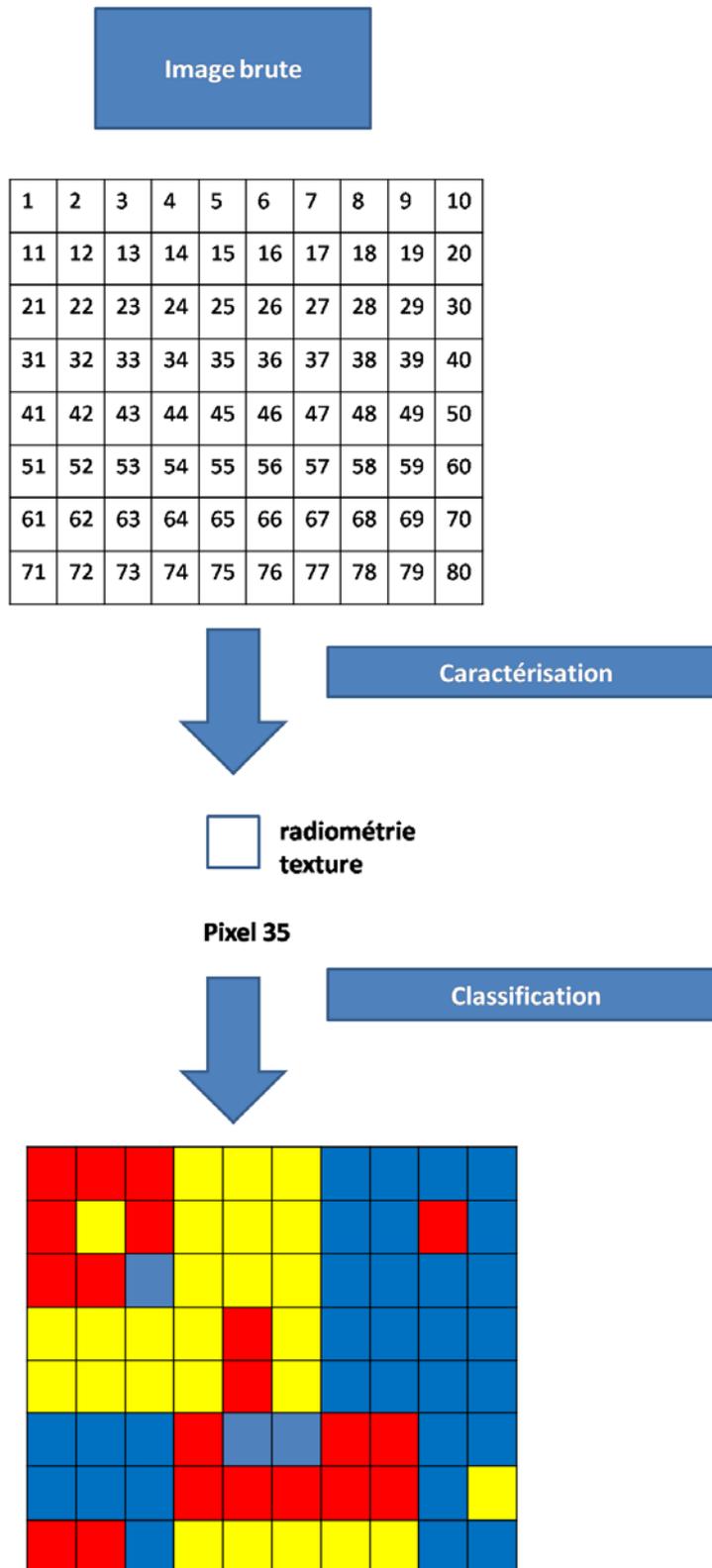


Figure 1.3 — Exemple d'analyse basée pixels sur une image artificielle.

Pour la caractérisation et le suivi de l'extension du tissu urbain de la région d'Alger, Smara *et al.* [Smara *et al.*, 2003] présentent des techniques de fusion et de classification floue d'images multi-sources optiques et radar. La méthodologie développée se base sur un processus comprenant différentes phases dont la première concerne les prétraitements de toutes les données mises en jeu (superposition géométrique, réduction du speckle dans les images radar SAR, ...).

1.2.2 Classification

1.2.2.1 Définition

La classification est le processus qui permet d'affecter une étiquette ou un label à un ensemble d'individus en se basant sur des propriétés qui les caractérisent. Pour l'analyse d'images orientée pixels, la classification est donc une partition de l'ensemble des pixels de l'image en un ensemble de classes. Nous distinguons deux familles de techniques de classification d'images : supervisées et non supervisées.

1.2.2.2 Classification supervisée

Il s'agit d'une technique très courante fondée sur l'analyse statistique d'échantillons. Cette classification demande à l'utilisateur d'instruire le système en désignant des zones de l'image comme étant des échantillons représentatifs des classes à extraire. La classification est précédée d'un apprentissage. Les fondements de l'apprentissage automatique sont les mathématiques et la théorie des probabilités.

Une étape très importante dans le processus de classification est celle du choix du classifieur. En effet, le choix du classifieur pèse dans le résultat de la classification. Parmi les classifieurs les plus utilisés nous pouvons citer le "classifieur par le maximum de vraisemblance" et le "classifieur par la distance minimale" [Duda et Hart, 1973, Mather, 1987]. Nous pouvons aussi citer le "classifieur par arbre de décision binaire" [Middelkoop *et al.*, 1989], le "classifieur par K plus proches voisins" [Dasarathy, 1990, Bezdek *et al.*, 1999], les "C-moyennes nettes" [Cover et Hart, 1967] et le classifieur bayésien sont aussi d'autres exemples de classifieurs classiques et très évoqués dans la littérature.

1.2.2.3 Classification non supervisée

La classification non supervisée ne demande aucune connaissance a priori. Cette méthode est souvent utilisée dans le cas où aucune information préliminaire n'est disponible sur les objets à classer. Deux cas peuvent se présenter : (i) il existe déjà des classes qui sont définies indépendamment de l'image et donc sans passer par un apprentissage, (ii) le nombre de classes et les caractéristiques qui leur sont associées sont définies automatiquement lors de la classification elle-même [Forgy, 1965].

Le principe de la classification non supervisée [Diday *et al.*, 1982, Jain *et al.*, 1999] consiste donc à regrouper les pixels selon la similarité de leurs caractéristiques, à déterminer la signature de chacune des classes constituées et à en rechercher la signification.

1.3 Du pixel vers l'objet

Malgré l'apport indéniable des approches d'analyse d'images basées pixels pour les images de télédétection à basse et moyenne résolution, ces approches ne sont pas adaptées à la haute et la très haute résolution [Blaschke et Strobel, 2001, Hay et Castilla, 2006]. Avec la complexité et la diversité présentes dans ces images, des approches se rapprochant de l'interprétation humaine sont de plus en plus utilisées. En effet, lorsque nous utilisons une machine pour classer une image, nous lui transmettons nos connaissances et nous essayons d'imiter certaines caractéristiques de l'interprétation humaine des images. À cet égard, et en particulier lors du passage au-delà des indices spectraux de certains objets géographiques, une approche de classification basée pixels devient limitée. La différence entre l'interprétation humaine des images et la modélisation pixélique de l'image devient évidente en remarquant que l'oeil humain voit plus qu'un ensemble de différentes couleurs [Lang *et al.*, 2009]. En effet, elle perçoit aussi des formes, des textures et des arrangements spatiaux des objets formant l'image. L'utilisation unique de la radiométrie ou de la texture devient insuffisante et tributaire d'une homogénéité radiométrique des objets. N'oublions pas les cas où des classes différentes peuvent avoir les mêmes propriétés radiométriques telles que les routes et les toits de quelques bâtiments. L'utilisation de l'unique infor-

mation radiométrique ne permet plus de répondre à ces attentes et le besoin de plus de propriétés devient nécessaire.

La figure 1.4 montre l'importance du contexte dans l'analyse d'une image et l'apport de l'arrangement spatial pour l'analyse finale. En regardant la première image en haut à gauche, un interprète humain décrira l'image comme étant une ligne blanche dans un fond rougeâtre. Avec une meilleure mise en contexte aperçue sur la deuxième image de haut à droite, l'avis de l'interprète risque de changer et de se pencher plus vers la peau d'un zèbre. L'image en bas à gauche nous conduira à une interprétation finale, c'est une chaise décorée avec un tissu de peau de zèbre. Plus il y a d'information contextuelle et d'arrangement spatial et plus l'interprétation est précise [Lang *et al.*, 2009].

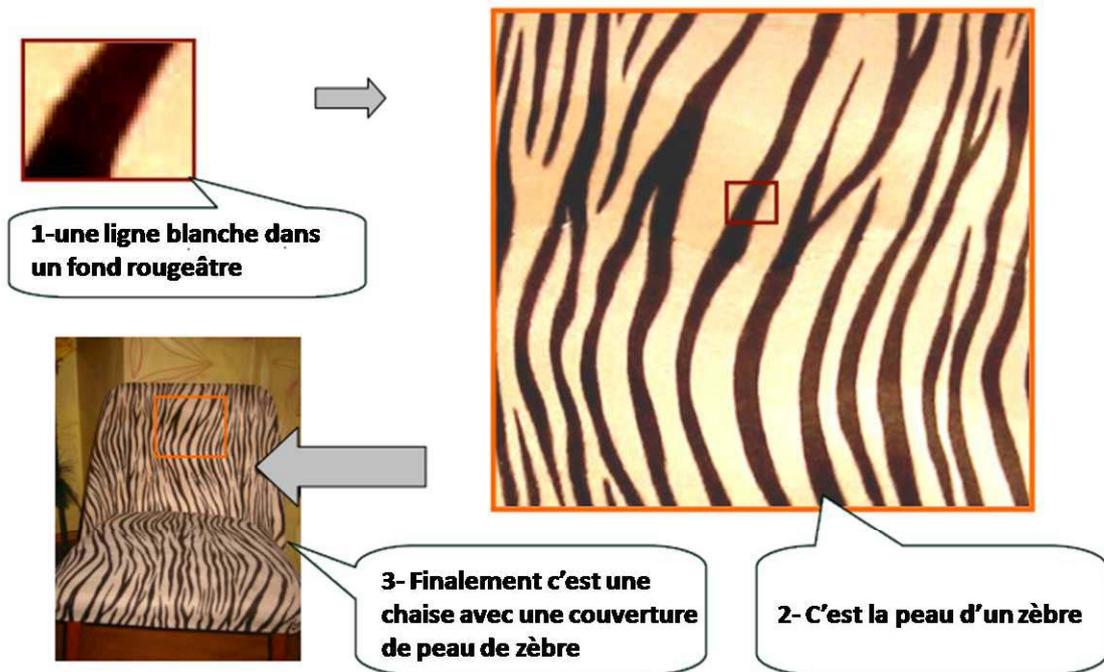


Figure 1.4 — Illustration de l'apport de l'information contextuelle dans l'analyse des images [Lang *et al.*, 2009]

Nous pouvons conclure que l'analyse d'image basée pixels souffre de certaines limitations qui font qu'elle ne donne plus satisfaction lorsque nous passons à la très haute résolution. Sa vision locale focalisée sur le pixel et ne prenant pas en compte son voisinage et son contexte, la pauvreté en attributs pouvant décrire le pixel qui se limite, la majorité du temps, à la radiométrie et à la texture et ignorant des attributs très influents comme lors du processus de décision (e.g., la forme, le voisinage, le contexte) ainsi que son manque d'adaptation à la complexité croissante des images, ainsi que les détails de plus en plus présents dans les images font que ces approches sont très rarement utilisées en analyse d'images satellitaires à très haute résolution.

Ces limites ont poussé les recherches vers les approches d'analyse d'images basées objets [Blaschke, 2010] qui représentent une alternative efficace permettant une vue "objet" de l'image. En effet, les approches objets permettent de considérer les caractéristiques spatiales des objets géographiques telles que la forme [Hay et Castilla, 2006, Kurtz, 2012]. Ces approches s'inspirent des principes de la vision humaine et de la manière avec laquelle l'être humain interprète une image. Le principe de base de ces approches est de diviser l'image en un ensemble de segments homogènes en utilisant un processus de segmentation. Or, plusieurs études ont prouvé qu'une étape de segmentation est la première étape d'un processus de perception visuelle [Zhaoping, 2000]. De plus, ils ont prouvé que cette première étape est totalement indépendante d'un quelconque apprentissage ou d'une connaissance a priori sur le monde observé [Roelfsema *et al.*, 1998]. Il en est de même pour la segmentation utilisée par les approches objets et qui se basent généralement sur les propriétés bas niveau de l'image.

Durant la dernière décennie, et afin de valider ces conclusions, plusieurs auteurs se sont intéressés à la comparaison des approches objets par rapport aux approches pixels [QIAN *et al.*, 2007, Matinfar *et al.*, 2007, Gao et Mas, 2008, Dehavri et Heck, 2009, Gholoobi *et al.*, 2010, Robert *et al.*, 2010, Myint *et al.*, 2011] .

Mantifar *et al.* [Matinfar *et al.*, 2007] comparent les résultats obtenus après l'application des approches pixels et objets sur une image LANDSAT7 ayant six bandes. Les auteurs ont utilisé une classification supervisée en utilisant la distance minimale alors que pour l'approche objets, ils ont utilisé l'outil "eCognition"¹ pour l'étape de segmentation et de classification. Pour la segmentation, ils ont appliqué la segmenta-

1. Copyright © 2013 Trimble/ www.ecognition.com

tion multi-résolution offerte par "eCognition". Ensuite ils ont appliqué la classification par plus proche voisin. Les résultats obtenus pour la moyenne de bonne classification basés sur la vérité terrain ont été de 81% pour la classification pixels contre 91% pour la classification objets ce qui prouve la pertinence des approches objets par rapport aux approches pixels.

Robert *et al.* [Robert *et al.*, 2010] comparent les approches pixels et objets pour la classification des couvertures terrestres. Ils ont démontré que les approches pixels présentent de faibles résultats de moyenne de bonne classification par rapport aux approches objets. Ceci est dû au fait que les approches pixels sont fondées sur les informations spectrales sans prendre en compte les informations texturales et contextuelles des pixels traités. Les approches pixels produisent aussi généralement un effet poivre et sel qui conduit à une mauvaise classification. Afin de confirmer ces conclusions, les auteurs ont comparé trois méthodes ; basée objets, supervisée et non supervisée en utilisant une image de la zone rurale de Hot Spring en Arkansas. Ils ont aussi utilisé plusieurs images afin de vérifier la robustesse des méthodes testées par rapport au changement du temps de l'acquisition de l'image, en utilisant deux images SPOT-5, une prise au printemps et l'autre en hiver et par rapport à la résolution, en utilisant une image aérienne CIR d'une résolution de 1m. En combinant ces images, les résultats montrent que les approches objets donnent de meilleurs résultats par rapport aux approches pixels (supervisées ou non supervisées) qui produisent des résultats fragmentés.

Myint *et al.* [Myint *et al.*, 2011] ont montré l'apport de l'analyse d'images basée objets par rapport aux approches pixels. Pour ce faire, les auteurs ont appliqué une première classification basée pixels en utilisant une approche d'analyse discriminante linéaire sur une image Quickbird à très haute résolution de la zone urbaine de la ville de Phoenix en Arizona. Pour les approches objets, les auteurs ont commencé par une première segmentation multi-résolution basée sur une approche d'évolution fractale en utilisant l'outil Definiens Developer 7.0. Pour l'étape de classification, les auteurs ont utilisé deux méthodes qui sont le classifieur par fonctions d'appartenance qui est fondé sur les connaissances expertes et le classifieur par plus proches voisins. Les résultats décrivant les taux moyens de bonne classification obtenus par le classifieur pixélique varient entre 50% et 67% selon la classe traitée alors que les résultats des deux classifieurs de l'approche objets varient entre 81% et 99% selon la classe traitée. Les auteurs ont conclu que les approches pixéliques ne sont pas très adaptées et efficaces

pour l'identification des classes des zones urbaines, ceci est dû à la résolution de l'image et aux différents détails qu'elle contient.

Devi et Krishna [Devi et Krishna, 2012] présentent une comparaison des approches pixels et objets en étudiant la classification de la couverture du sol d'une image LISS 4 de la région de Nalgonda à Andhra en Inde. La figure 1.5 présente les étapes suivies pour la comparaison. Pour la classification pixels, les auteurs utilisent un classifieur par maximum de vraisemblance. Ensuite, ils génèrent une matrice d'erreurs permettant d'évaluer cette classification qui a permis d'obtenir un taux de classification moyen avoisinant les 78%. Pour la classification objets, les auteurs commencent par une segmentation multi-résolution utilisant la version 8 du logiciel "eCognition". Ensuite, ils appliquent une classification par plus proches voisins qui permet d'obtenir un taux de classification moyen avoisinant les 85% prouvant l'amélioration des résultats des approches objets par rapport aux approches pixels.

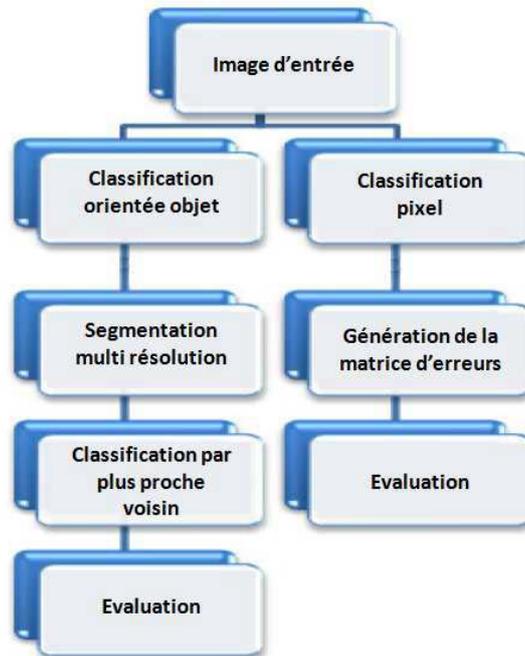


Figure 1.5 — Workflow de la comparaison des approches objet et pixel.

Pour résumer, les limites des approches d'analyse d'image basées pixels ont été démontrées et l'apport des approches objets, par rapport aux approches pixéliques, est indéniable dans le cadre de l'analyse d'images à très haute résolution spatiale. Dans la section suivante, nous présentons les principes des approches d'analyse d'image basée objets ainsi qu'une revue des approches présentées dans la littérature.

1.4 Analyse d'images basée objets

L'analyse d'images basée objets a été définie lors de la conférence internationale de l'analyse d'image basée objets OBIA'2006 (en référence à son appellation anglo-saxonne "Object Based Image Analysis") comme : "une sous discipline de la science de l'information géographique consacrée à subdiviser les images de télédétection en des objets image significatifs (i.e. zones d'intérêt), et à évaluer leurs caractéristiques selon des échelles spatiales, spectrales et temporelles" [Hay et Castilla, 2006, Sebari, 2008].

1.4.1 Principe des approches basées objets

L'analyse d'images basée objets est une alternative proposée pour palier aux limites des approches orientées pixels [Blaschke, 2010]. Le principe de base de ces approches est le passage d'une vision pixélique de l'image vers une vision objets [Zhou et Troy, 2008, Dennis *et al.*, 2012]. L'unité de base du traitement se transforme donc en un ensemble de pixels homogènes regroupés dans une même région (i.e. objet). L'étape d'extraction de ces objets est l'étape initiale de toute approche objets. Il est généralement convenu que l'analyse d'images basée objets est fondée sur plusieurs anciens concepts utilisés dans la télédétection [Sebari, 2008] tels que la segmentation, la détection de contour, l'extraction des propriétés des images et la classification. La figure 1.6 illustre les trois principales étapes de l'analyse d'images basée objets.

Le passage du pixel vers l'objet est l'étape initiale des approches objets. Elle est souvent réalisée par une méthode de segmentation. Le critère d'homogénéité de la segmentation est préalablement défini selon une propriété possédant une mesure de similarité. Les critères d'homogénéité les plus utilisés sont la radiométrie et la texture.

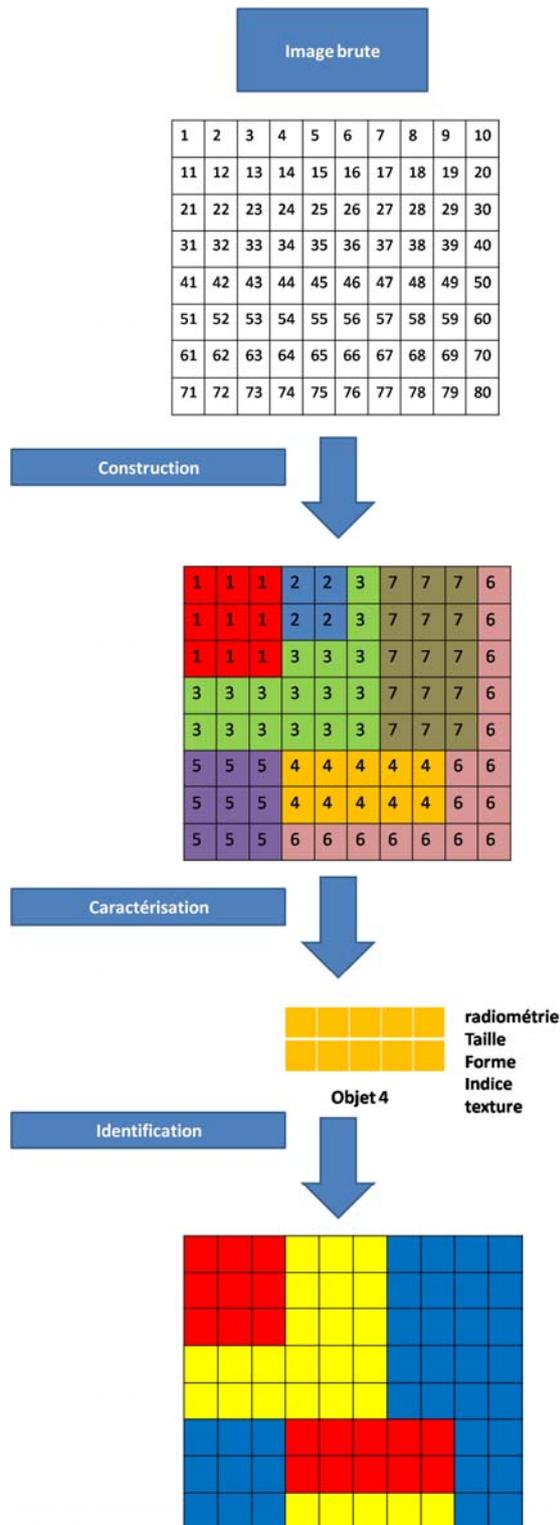


Figure 1.6 — Exemple d'analyse d'image basée objets sur une image artificielle

Une fois les objets extraits, ils peuvent être décrits à travers plusieurs caractéristiques. Le passage à l'objet permet d'élargir l'ensemble des caractéristiques utilisées, rendant notamment possible l'utilisation de propriétés géométriques et contextuelles. La prise en compte de ces diverses informations permet d'améliorer les résultats [Blaschke et Lang, 2006]. En effet, la ligne directrice principale des approches objets est de se rapprocher le plus possible de la manière dont le cerveau humain analyse une image. Elle vise à exploiter d'une façon appropriée les données de télédétection en se basant sur des objets image ressemblant aux objets réels au lieu d'employer les pixels comme base pour la classification. La vision humaine ne percevant pas une image comme un ensemble de pixels, mais tend d'abord à diviser l'image en zones homogènes, qui sont plus facilement caractérisables par des informations pertinentes telles que l'écart-type de la luminance des pixels constituant l'objet, sa forme géométrique, ainsi que d'autres propriétés [Sebari, 2008, Blaschke, 2010]. La dernière étape est l'identification des objets en se basant sur les propriétés extraites. Passant du pixel vers l'objet, cette convergence permet de traiter directement les objets d'intérêt présents dans l'image. La notion de sémantique prend alors toute son ampleur en passant d'une vision pixélique visant uniquement l'information radiométrique vers une vision sémantique et utilisant plusieurs informations haut niveau [Forestier *et al.*, 2008a, Derivaux *et al.*, 2010].

1.4.2 Segmentation d'images

La segmentation consiste à créer des partitions dans l'image en identifiant des pixels similaires. Le critère de similarité permet de contrôler l'aspect final de la segmentation. La mesure de la similarité de deux pixels est donc le point clé des techniques de segmentation. Il convient donc de choisir avec soin la méthode de calcul de la similarité. Nous énumérons plusieurs méthodes de regroupement des pixels tels que le regroupement par couleur, le regroupement par homogénéité, ou encore le regroupement par taille [Pal et Pal, 1993].

La segmentation a pour objectif de différencier des zones d'intérêt dans l'image. C'est généralement une première étape d'un traitement plus complexe comme la reconnaissance de formes. La segmentation est pourtant un problème complexe et mal posé : complexe car pouvant générer de nombreuses solutions et mal posé car il est difficile de savoir si une solution est meilleure qu'une autre [Derivaux, 2009]. Nous présentons

une classification de ces approches en quatre grandes familles :

- segmentation par pixel (par seuillage) ;
- segmentation par contour ;
- segmentation par région ;
- segmentation par collaboration région/contour.

1.4.2.1 Segmentation par pixel (par seuillage)

Ces techniques se basent essentiellement sur la radiométrie des pixels. Elle repose sur l'analyse de l'histogramme de l'image [Pal et Pal, 1993]. Nous partons d'un rapport qu'entretient chaque pixel individuellement avec des informations calculées sur toute l'image, comme par exemple la moyenne des niveaux de gris de l'ensemble des pixels, ou la médiane, permettant de construire n classes d'intensité. La détection des régions se fait en détectant un ou plusieurs seuils permettant la meilleure séparation entre les classes (régions) seuil, on parle de seuillage. Les pixels appartenant à une même classe et étant connexes forment des régions. Les limitations de ces techniques consistent dans le fait qu'elles ne prennent pas en considération les informations spatiales de l'image et ne considèrent pas le pixel dans son voisinage. Leur sensibilité au bruit fait qu'elles ne sont pas adaptées aux images à très haute résolution qui sont caractérisées par une quantité énorme d'informations et de détails [Plataniotis et Venetsanopoulos, 2000, Sebari, 2008].

1.4.2.2 Segmentation par contour

Ces techniques cherchent à exploiter le fait qu'il existe une transition détectable entre deux régions connexes. Plusieurs de ces techniques sont fondées sur l'application d'opérateurs de détection de contours tels que les opérateurs de Canny [Canny, 1986] ou Dérivée [Deriche, 1987]. Le problème de ces techniques est qu'elles peuvent présenter des régions morcelées et peu précises, les contours n'étant pas fermés, une étape de post-traitement est nécessaire pour les raccorder. L'une des techniques les plus connues de segmentation par contour est la ligne de partage des eaux [Vincent et Soille, 1991].

1.4.2.3 Segmentation par région

La segmentation par région consiste à construire des surfaces en regroupant des pixels voisins suivant un critère d'homogénéité. Les techniques à base de crois-

sance de régions figurent parmi les plus utilisées dans le cas d'images satellitaires [Kurtz, 2012]. Les techniques les plus connues sont les techniques de croissance de région [Ericksson, 2004] et de division fusion [Yang et Lee, 1997]. Malgré le fait qu'elles présentent des segments connectés, ceux-ci ne coïncident pas nécessairement avec les limites des objets d'intérêt.

1.4.2.4 Segmentation par collaboration région/contour

Ces approches se basent sur la complémentarité entre les approches contours et régions. En effet, les approches contours présentent la limite de ne pas avoir de contours fermés alors que les approches régions ne coïncident pas avec les contours de l'image. Plusieurs types de coopérations ont été proposés dans la littérature [Sebari, 2008] :

- Coopération séquentielle où l'une des deux techniques est réalisée en premier lieu. Le résultat obtenu est alors utilisé comme contrainte pour l'autre technique.
- Fusion des résultats obtenus par l'application indépendante d'un processus de segmentation région et d'un processus de segmentation contour.
- Coopération mutuelle où il y a une interaction entre les deux processus de segmentation qui vont s'échanger leurs informations, permettant ainsi une meilleure segmentation.

Une fois les objets extraits, vient l'étape de caractérisation des objets par un ensemble d'attributs permettant la description de ces derniers.

1.4.3 Caractérisation des objets

La caractérisation des objets est une étape primordiale pour l'analyse des images orientée objets. Le passage du pixel vers l'objet permet une meilleure caractérisation de ces derniers et offre une plus large palette d'attributs les caractérisant.

- Les attributs spectraux sont dérivés des données brutes de l'image. Ils se basent sur la réponse spectrale des pixels dans les différentes bandes de l'image. Lors du passage à l'unité objet, ces attributs peuvent être perçus globalement en extrayant pour chaque objet traité la moyenne, l'écart-type, la valeur minimale ou maximale des valeurs radiométriques associées aux pixels qui le composent. Cette vision objet permet de fournir une information plus pertinente et discriminante sur la composition spectrale de l'objet (i.e., région) et de synthétiser les informations

radiométriques attachées aux pixels.

- Les attributs texturaux : Haralick et Gagallowicz [Haralick, 1979] définissent la texture comme un phénomène à deux niveaux, le premier concerne la description des primitives dont est composée l'image et qui sont ses propriétés caractéristiques, le deuxième implique la dépendance spatiale entre ces primitives [Haralick, 1979].

Les caractéristiques de Haralick sont l'une des caractéristiques de texture les plus utilisées, nous en citons la moyenne, la variance, l'énergie, l'entropie, le contraste, l'homogénéité, ...

- Les attributs géométriques se focalisent essentiellement sur les attributs de forme et de taille. Les indices de formes sont très utilisés pour caractériser la forme des objets de l'image [ROSIN, 1999, Bendjoudi, 2002, Puissant, 2003, Thibault *et al.*, 2007, Sebari, 2008, Sellaouti, 2009, Hentati *et al.*, 2011].
- Les attributs contextuels permettent de caractériser les informations spatiales entre les objets. Les relations spatiales entre les objets sont des informations très intéressantes car elles sont plus stables que les propriétés des objets elles-mêmes. Elles fournissent une information structurelle et peuvent constituer un outil de description des objets et un outil qui permettra d'aider à la décision en cas d'ambiguïté. Elles représentent un moyen efficace de décrire des scènes structurées et sont potentiellement utiles dans de nombreux domaines d'application [Bloch, 2002, Colliot *et al.*, 2005, Inglada et Michel, 2009]. Nous pouvons distinguer différentes relations spatiales telles que l'adjacence, l'inclusion, la distance entre les objets, la direction, ...

1.4.4 Identification des objets

Le passage du pixel vers l'objet a changé la vision du problème d'identification des objets. Les informations supplémentaires et la nouvelle possibilité de faire le parallèle entre les objets d'intérêt présents dans l'image et les objets extraits ont ouvert de nouvelles perspectives. En effet, de nouvelles techniques d'interprétations qui ne pouvaient pas être utilisées avec les pixels ont vu le jour [Kurtz, 2012]. [Sebari, 2008] a classé les méthodes d'identification en deux familles :

- les méthodes par attributs,
- les méthodes basées connaissances.

1.4.4.1 Les méthodes par attributs

Ces méthodes sont les méthodes classiques de classification auxquelles les nouveaux attributs objets sont intégrés pour une meilleure discrimination entre les classes composant l'image [Eriksson, 2005, Forestier *et al.*, 2008b, Dhiaf et Sellaouti, 2009, Gao *et al.*, 2011]. En effet, la nouvelle palette d'attributs permet de compléter les simples informations radiométriques qui caractérisaient les pixels de l'image, généralement employées dans les approches basées pixels. Ces attributs radiométriques sont insuffisants lorsque les objets d'intérêts appartenant à des classes distinctes ont des valeurs radiométriques similaires, ce qui conduit à l'introduction de nouvelles caractéristiques qui sont plus discriminantes permettant une meilleure séparation de ces objets. L'exemple de certains bâtiments et route qui ont une signature radiométrique très proche mais des propriétés géométriques distinctes montre l'apport de ces informations.

1.4.4.2 Les méthodes basées connaissances

Ces méthodes se fondent sur le fait que les approches objets facilitent le mécanisme d'intégration de connaissances haut niveau. Plusieurs modélisations des connaissances ont été proposées. Les ontologies ont été introduites pour une formalisation des classes en hiérarchie de concepts [Maillot *et al.*, 2003, Maillot, 2005, Derivaux *et al.*, 2007, Forestier *et al.*, 2012]. Les systèmes à base de règles ont été aussi utilisés afin d'identifier des objets d'intérêts.

Suzuki [Suzuki, 2002] modélise les connaissances avec un système d'inférence flou. En effet, la connaissance est représentée en deux niveaux. Le premier niveau est la représentation sémantique de la connaissance, chaque classe est décrite par un ensemble appelé conclusion qui représente la description sémantique de la classe. Les conclusions sont ensuite décomposées en un ensemble de variables mesurables des sous ensembles flous.

Des détecteurs spécifiques aux classes extraits à partir de la base de connaissance ont été utilisés par Derivaux [Derivaux, 2009]. Chaque détecteur commence par considérer que toute l'image correspond à la classe qu'il recherche. Ensuite, une succession de filtres est appliquée afin de retirer les zones de l'image qui ne correspondent pas à la classe recherchée. Une fois ces filtres appliqués, il ne reste plus que les zones de l'image

pouvant appartenir à la classe d'intérêt.

Les graphes sont aussi utilisés pour modéliser les connaissances [Deruyver *et al.*, 2011]. En effet, l'image est modélisée par un graphe où chaque région (i.e., objet) de la scène à analyser est représentée par un nœud dans le graphe de donnée. Un appariement entre le graphe représentant l'image et un graphe de référence porteur de connaissances expertes permet d'identifier les objets de l'image. Les arcs du graphe sont utilisés pour modéliser les connaissances spatiales. Les autres connaissances pourront être les attributs qui décorent les arcs ainsi que les nœuds. Deruyver et Hodé [Deruyver et Hodé, 1997] modélisent l'ensemble des connaissances concernant le cerveau humain sous forme d'un graphe aléatoire comme illustré dans la figure 1.7.

1.5 Etat de l'art et positionnement

Plusieurs approches d'analyse d'images basée objets ont été proposées dans la littérature lors des deux dernières décennies. Blaschke en a présenté en 2010 un état de l'art très complet [Blaschke, 2010]. En 2009, près de 800 articles scientifiques visant des travaux sur les approches d'analyse d'image orientée objets ont été recensées dans la littérature. Ce nombre dépasse fort probablement les 1000 de nos jours, ce qui prouve l'importance que prend cette famille d'approches dans l'analyse d'images satellitaires et spécialement l'analyse des images à très haute résolution [Giraudon *et al.*, 1992, Blaschke et Strobel, 2001, Marangoz *et al.*, 2004, Durand *et al.*, 2007, Karsenty *et al.*, 2007, Jacquin *et al.*, 2008, Forestier *et al.*, 2008c, Aminipouri *et al.*, 2009, Salehi *et al.*, 2011, Aksoy et Ercanoglu, 2012, Attarzadeh et Momeni, 2012, Grigillo et Kanjir, 2012, Seema, 2012, Mahmoudi *et al.*, 2013]. Dans cette section, nous introduisons une classification des approches d'analyse d'images basées objets en 3 familles.

1.5.1 Approche naturelle

L'approche naturelle repose sur la structure classique de l'analyse d'image orientée objets qui est composée d'une phase de segmentation de l'image en un ensemble d'objets. Ces objets sont décrits par un ensemble d'attributs, puis classés.

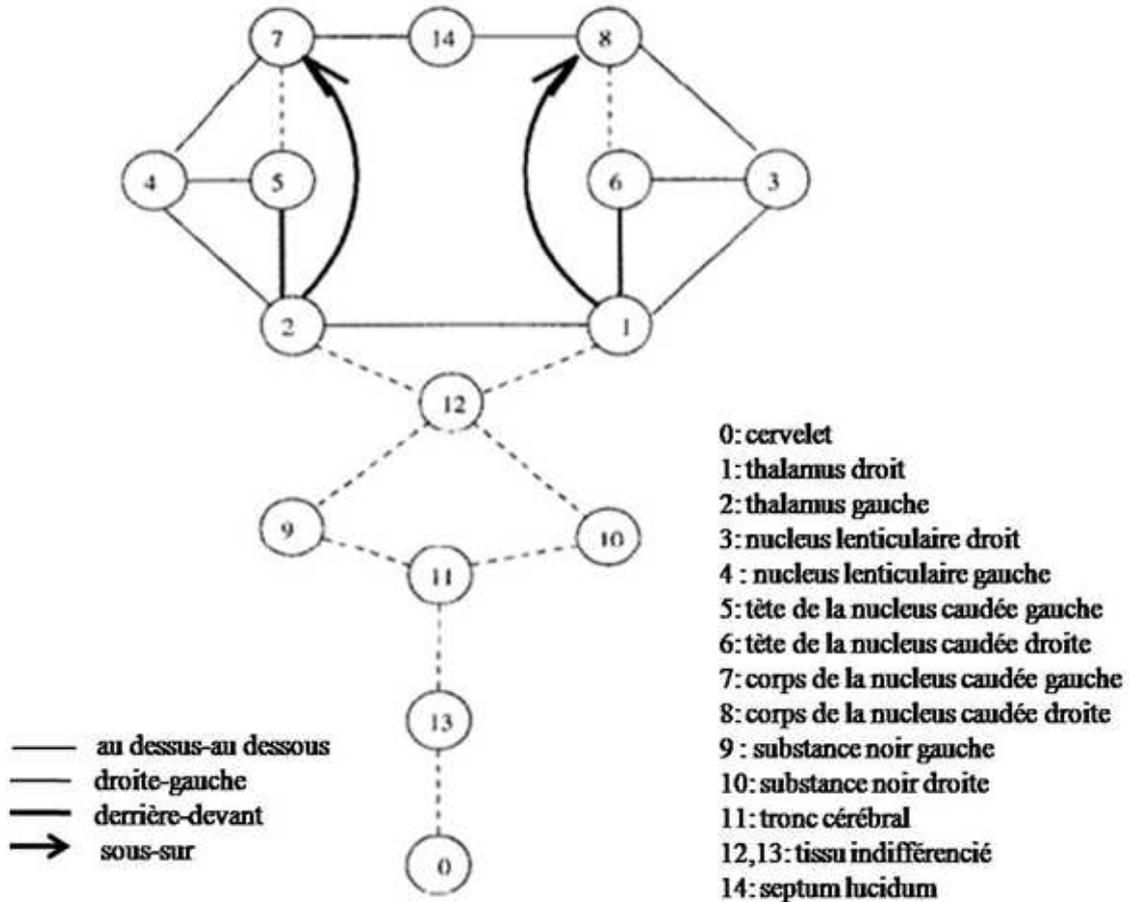


Figure 1.7 — Graphe de référence du cerveau humain.

[Walter, 2004] présente une approche pour la détection des changements apparus dans les zones urbaines. Cette approche est composée en deux parties, la première est une analyse d'images basée objets dont la phase d'identification consiste en une classification supervisée avec l'algorithme de maximum de vraisemblance. La base d'apprentissage est extraite à partir d'une base GIS existante. Dans la deuxième partie, une mise en correspondance entre les objets classés et les objets existant dans la base GIS est appliquée pour détecter les changements. L'auteur utilise toutes les bandes spectrales. Les caractéristiques utilisées sont des caractéristiques spectrales et texturales. La figure 1.8 illustre les différentes étapes utilisées.

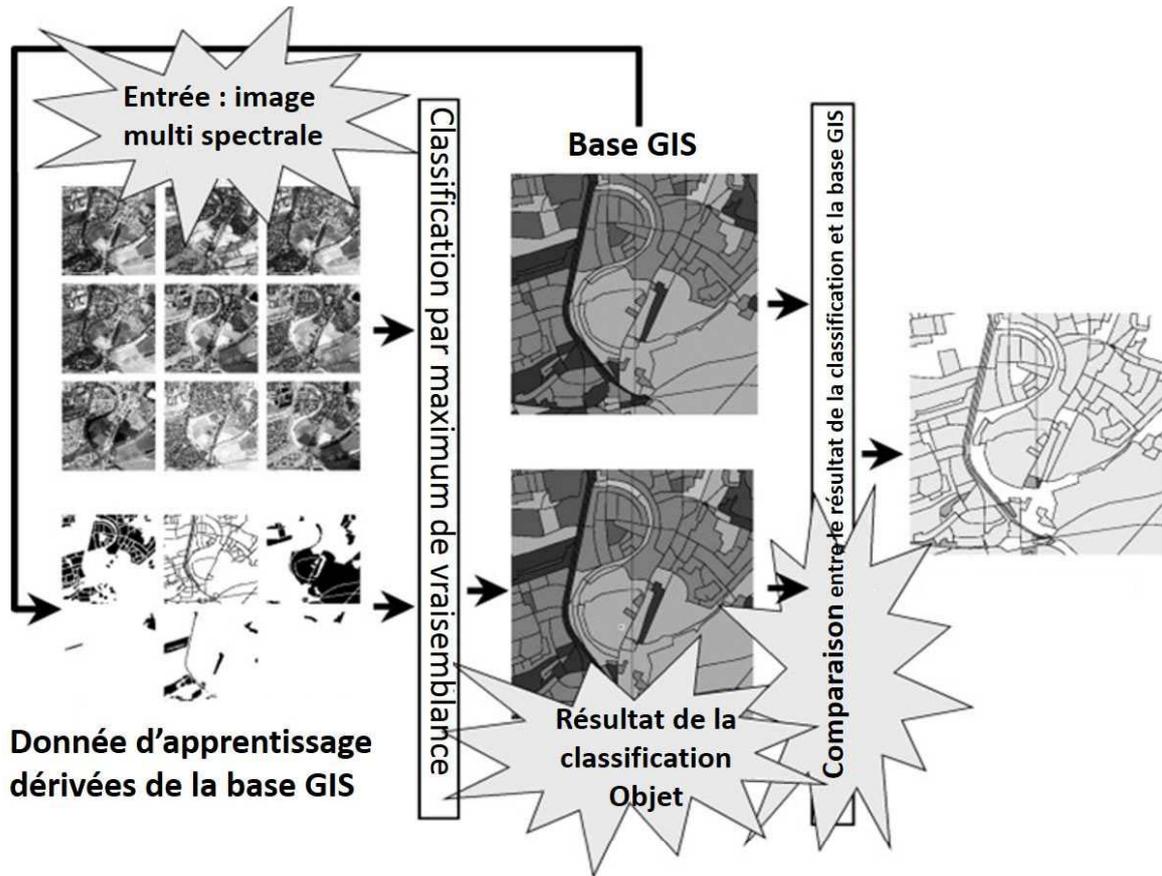


Figure 1.8 — Workflow de l'approche naturelle proposée par Walter [Walter, 2004].

[Ericksson, 2004] utilise une approche basée objets pour détecter les couronnes d'arbres. En effet, il commence par une segmentation qui va lui permettre de détecter les couronnes d'arbres. Cette segmentation est la croissance de région par mouvement Brownien. Une fois les couronnes des arbres détectées, Ericksson procède à une classification de ces couronnes en se basant sur un ensemble de règles spécifiques à chaque classe. La couronne qui vérifie les règles d'une classe est affectée à cette dernière.

[Dhiaf et Sellaouti, 2009], utilisent les couronnes extraites par l'algorithme de croissance de région pour procéder à une classification basée sur les indices de formes et la texture.

[Lefebvre *et al.*, 2011] présentent une analyse d'images fondée sur les ondelettes et la théorie des évidences. Une première étape de segmentation utilisant la Ligne de Partage des Eaux (LPE) est appliquée sur une image de contours créée à partir des composantes horizontales et verticales de la décomposition en ondelettes afin d'éviter

un résultat sur-segmenté. Ensuite, une agrégation des petits objets non exploitables dans l'étape suivante est effectuée en fusionnant les objets voisins de petite taille et de valeur moyenne de luminance proche. La deuxième étape est la caractérisation des objets extraits avec des attributs de luminance et de texture. La dernière étape est la classification basée sur la théorie de l'évidence en fusionnant les critères de similarité de luminance et de texture.

La majorité de ces approches se base sur l'hypothèse que les objets extraits lors de la phase de segmentation représentent des objets d'intérêt ce qui permet de faire directement le parallèle entre ces objets et les objets d'intérêts.

1.5.2 Approche hiérarchique

Cette approche utilise une représentation multi-échelle de l'image. En effet, elle repose sur une segmentation multi-échelle, l'image est représentée par une famille d'images allant de la perception de l'image la plus fine vers la plus grossière. La segmentation multi-résolution est une heuristique d'optimisation qui cherche à minimiser localement l'hétérogénéité moyenne des objets dans une résolution donnée. Cette segmentation reste tributaire des paramètres d'échelle, de couleurs et de forme qui sont difficiles à extraire sans l'utilisation d'un jeu d'essais-erreurs [Aksoy et Ercanoglu, 2012]. La figure 1.9 illustre la vision hiérarchique de l'image.

[Hofmann, 2001] présente une approche basée objets pour la classification des zones urbaines. Il introduit deux approches utilisées par le logiciel "eCognition" pour extraire les objets d'intérêt. Le départ de l'approche est une segmentation multi-échelle qui permet de créer un réseau hiérarchique représentant l'image. La segmentation multi-échelle peut être réalisée de deux façons différentes qui sont bottom-up et top-down. L'approche top-down commence par générer les objets du niveau le plus haut. Tous les objets des niveaux le plus bas sont des sous objets du niveau supérieur. L'approche bottom-up, quant à elle, opère inversement. La segmentation commence par générer les petits objets. Tous les objets générés dans les niveaux les plus hauts sont alors considérés comme des super objets du niveau initial.

Parmi les nombreux travaux utilisant cette approche, on peut citer

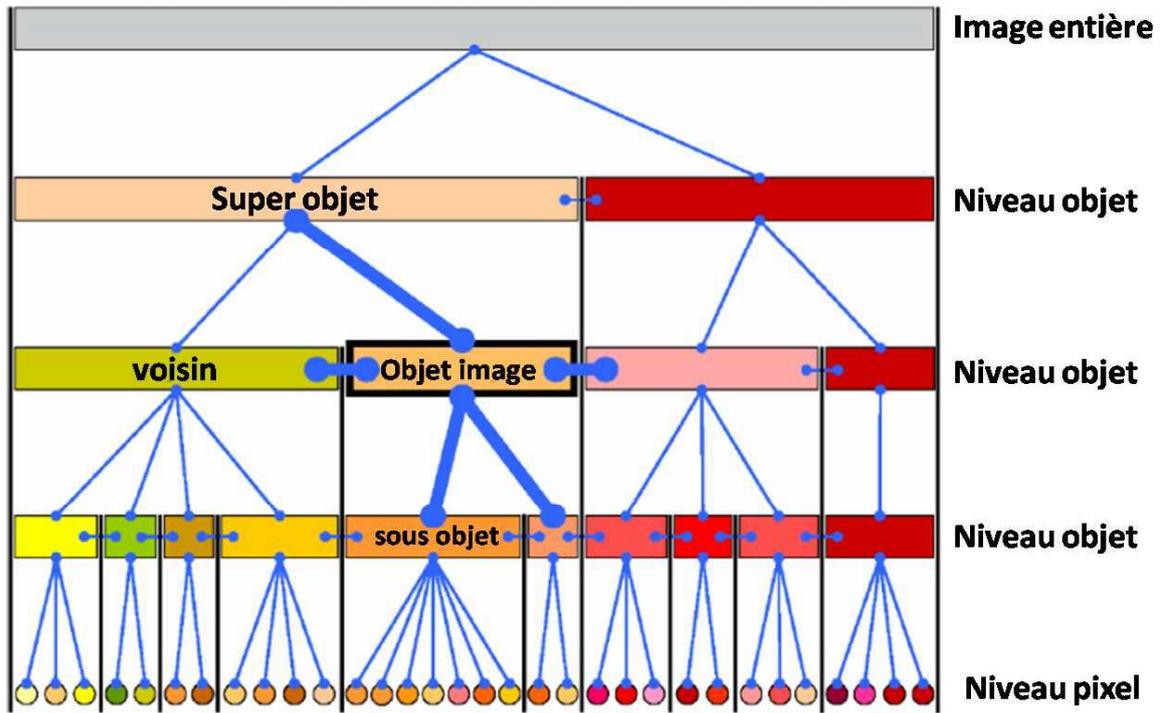


Figure 1.9 — Vision hiérarchique de l'image : chaque objet image est lié à son super objet et son sous objet.

[Marangoz *et al.*, 2004] qui l'utilise pour détecter les routes et les bâtiments dans une image IKONOS. [Karsenty *et al.*, 2007] utilisent la même approche pour évaluer la perméabilité des sols en zone urbaine à l'aide d'imagerie très haute résolution et de données laser scanner à Curitiba au Brésil. De même, pour [Jacquin *et al.*, 2008] qui utilisent cette approche pour voir l'effet de l'expansion urbaine sur les inondations et les anticiper.

Aksoy et Ercanoglu [Aksoy et Ercanoglu, 2012] présentent une approche basée objets permettant la classification des zones de glissement de terrain sur des images Landsat ETM+ et un modèle d'élévation de terrain. La figure 1.10 montre les différentes étapes de l'approche proposée. Toutes les étapes introduites ont été réalisées par le logiciel "Definiens Professional Earth" d'eCognition. L'un des problèmes rencontrés par les auteurs est la paramétrisation de la segmentation multi-résolution. La solution proposée est une solution expérimentale basée sur les essais-erreurs. Ensuite, un ensemble de dix caractéristiques de différentes natures a été extrait pour la caractérisation des objets extraits dans la phase de segmentation. Finalement, une classification floue a été utilisée pour la classification de ces objets. Tout le long de ce processus, des connaissances

et des informations supplémentaires ont été utilisées telles que la carte d'élévation ou la carte NDVI (Normalized Difference Vegetation Index).

Salehi *et al.* [Salehi *et al.*, 2011] présentent une approche hiérarchique pour la classification de la couverture terrestre profitant de la combinaison des approches pixels et objets. Une première segmentation multi-échelle est appliquée en utilisant la version 2010 du logiciel "eCognition" Developer. Trois paramètres qui sont l'échelle, la forme et la compacité doivent être introduits par l'utilisateur. Ces paramètres ont un impact direct sur le résultat final. Afin d'éviter le processus d'essais-erreurs pour trouver le meilleur paramétrage, les auteurs utilisent l'optimiseur de paramètres de segmentation flou qui permet de détecter le meilleur jeu de paramètres dans trois niveaux de la segmentation. Pour chaque échelle, un ensemble de caractéristiques spectrales, spatiales, texturales et morphologique sont extraits. Finalement un algorithme de classification par maximum de vraisemblance basée pixels a été appliqué à l'image originale ainsi qu'à l'ensemble des images créées à partir des attributs extraits pour classer l'image en cinq catégories différentes.

1.5.3 Approche hybride

Les approches hybrides sont des approches basées sur la cohabitation et la collaboration entre plusieurs approches objets pour l'identification des objets d'intérêts. [Giraudon *et al.*, 1992] ont développé le système MESSIE. Ce système est une architecture multi-spécialiste, bâti autour d'une architecture de type tableau noir, dont l'objectif est de réaliser une interprétation basée uniquement sur une connaissance générique des objets sans utiliser de connaissances exogènes à la donnée image. Les auteurs montrent comment une modélisation des objets physiques de la scène exprimée sous les quatre points de vue forme, contexte, aspect et fonction peut améliorer la classification. Ce système est capable de manipuler des connaissances ponctuelles (informations radiométriques) et structurelles (propriétés géométriques, relations spatiales). Chaque spécialiste a une tâche spécifique et indépendante des autres spécialistes. Pour communiquer entre eux, les spécialistes utilisent une zone mémoire commune qui est le tableau noir. Plusieurs rôles sont assignés aux spécialistes tels que l'extraction des objets, le calcul des caractéristiques structurelles, l'évaluation des hypothèses, l'étiquetage des objets, la détection des conflits, etc. Le système peut facilement accepter de nouvelles connaissances expertes et de nouveaux spécialistes.

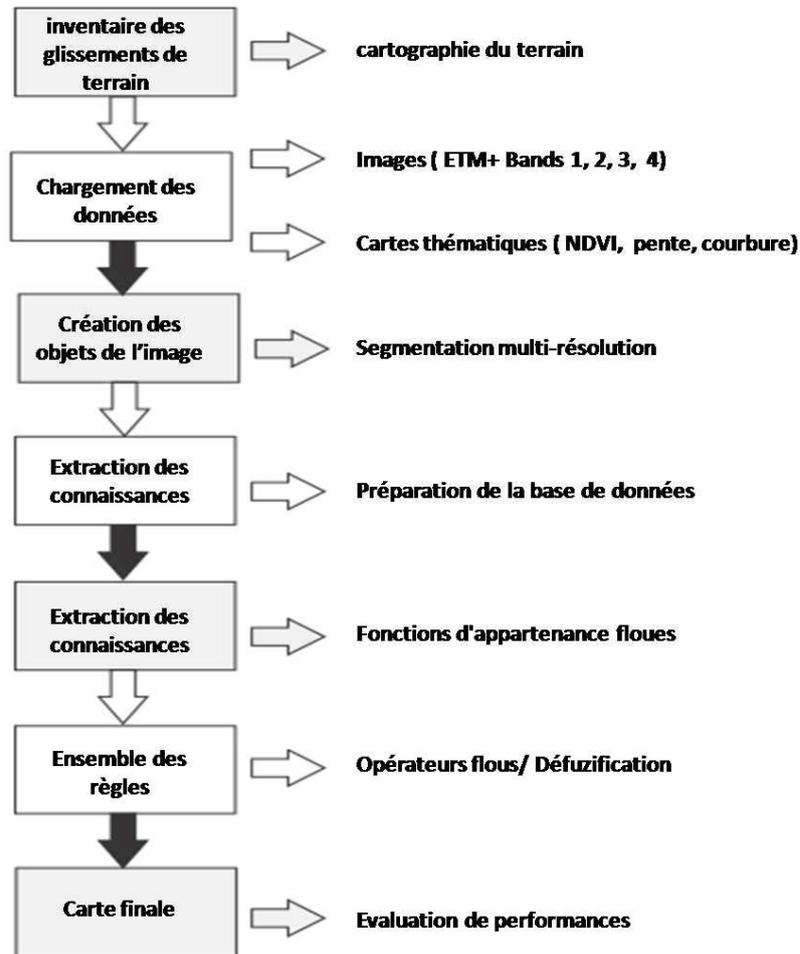


Figure 1.10 — Workflow de l'approche hiérarchique [Aksoy et Ercanoglu, 2012].

[Forestier *et al.*, 2008c], quant à eux, proposent une approche orientée objet collaborative et multi-stratégie. Elle intègre un ensemble de classifieurs non supervisés et présente une nouvelle approche qui permet de faire collaborer les différents classifieurs. L'originalité de cette approche est qu'elle intègre le processus de collaboration durant l'étape de classification. En effet, elle est divisée en trois étapes. La première est une classification initiale où chaque classifieur est lancé avec ses propres paramètres, permettant de créer des objets. La deuxième étape consiste en un raffinement des résultats divisés en deux parties à savoir d'une part l'évaluation de la similarité des classes et d'autre part, leur raffinement. La dernière étape est l'unification où les résultats raffinés sont unifiés avec un algorithme de vote.

Dans [Mahmoudi *et al.*, 2013], l'approche proposée est une approche pour l'extraction d'objets appliqué sur des images du satellite WorldView-2 et en utilisant un modèle numérique de terrain. Elle est composée de trois étapes : (i) une étape de prétraitement, (ii) une étape d'analyse d'image orientée objets et finalement, (iii) une étape de reconnaissance multi-agent. En effet, et après avoir procédé au prétraitement, les auteurs appliquent une segmentation suivie d'une classification pour labéliser les objets extraits de l'image. Afin d'améliorer les résultats obtenus après l'application de l'analyse d'image orientée objets, les auteurs utilisent les capacités des systèmes multi agents en appliquant des modifications dans les décisions prises dans la phase de classification. Pour la partie prétraitement, les auteurs utilisent un algorithme de détection d'occlusion basé sur le modèle numérique de terrain.

La première étape de la partie d'analyse d'images basée objets est une segmentation multi-résolution permettant d'extraire les objets de l'image. Ensuite, afin de classer les objets extraits, les auteurs utilisent un ensemble de descripteurs spectraux (indices de différence normalisée (NDI) et les ratios simples (SR)), texturaux, structurels et spatiaux. Après l'extraction des propriétés des objets, une classification fondée sur les connaissances est appliquée en se basant sur un ensemble de règles expertes. La classification proposée est une classification multi-processus. Dans la première étape, l'image entière est classée en fonction des règles de raisonnement spectrales. Ensuite, dans la deuxième étape, les objets sont réévalués en se basant sur des règles structurelles et texturales.

Afin d'améliorer la classification qui, malgré les résultats qu'elle présente, reste dépendante de la phase de segmentation, l'auteur propose un système multi-agents pour la reconnaissance des objets présents dans l'image. Ce système est composé de deux niveaux. Dans le premier niveau, les régions classées obtenues à partir de l'analyse d'images basée objets sont utilisées comme environnement des agents et l'objectif de chacun des agents de reconnaissance d'objets est de modifier les régions mal classées. Deux type d'agent sont définis pour chaque classe d'objets : un agent image qui analyse les propriétés texturales et spectrales des objets et un agent élévation qui analyse le modèle numérique de surface. Le rôle de ces agents est de détecter les régions mal classées. Après le partage d'informations collectées, l'agent coordinateur envoie des messages aux agents de chaque groupe de classes (image et élévation) afin de permettre à chaque agent de réviser les résultats de ces partenaires, de les analyser et

de les confronter. En cas de conflit, les agents utilisent un ensemble de règles permettant la résolution de ces derniers. Après cette phase, l'agent coordinateur envoie les différents objets non classés aux agents qui essaient de les reclasser. Tous les objets qui restent non classés seront traités dans la deuxième étape. Dans le deuxième niveau, les opérations sont quasi identiques au niveau 1 sauf que dans ce niveau, ce sont les propriétés structurelles et contextuelles qui sont utilisées.

Pour résumer, toutes les approches décrites ont une limite commune, c'est leur dépendance à la phase de segmentation. En effet, cette phase initiale est la plus importante vu qu'une mauvaise segmentation entraîne nécessairement une mauvaise classification. Dans toutes les approches précitées, l'étape de segmentation n'est jamais remise en cause par la classification sauf pour l'approche de [Mahmoudi *et al.*, 2013].

1.5.4 Limite des approches d'analyse d'images basée objets

Malgré l'apport des approches basées objets pour l'analyse d'images et les solutions qu'elles offrent par rapport aux limites des approches basées pixels, elles restent dépendantes des résultats obtenus dans la première étape de construction. En effet, ces approches se basent sur l'hypothèse que les objets construits lors de la phase de segmentation sont des objets du monde réel ce qui permet de faire le parallèle entre les objets d'intérêts et les objets construits dans la phase de segmentation. En effet, la méthode de segmentation doit permettre la création d'objets fermés et représentatifs des objets réels figurant dans l'image. Or, cette hypothèse n'est généralement pas correcte. En effet, comme déjà mentionné dans la section 1.4.2, la segmentation est de nature un problème mal posé et complexe, le résultat de la segmentation est donc dépendant de plusieurs paramètres et aucun algorithme de segmentation n'est générique. Ainsi, avec l'augmentation des détails dans les images de télédétection à très haute résolution, les objets d'intérêt deviennent de plus en plus complexes et hétérogènes, provoquant des problèmes lors de l'étape de segmentation surtout lorsque cette dernière ne prend pas en compte la nature de l'objet traité.

D'autre part, Gao a montré la dépendance entre le résultat de la segmentation et de la classification [Y. Gao *et al.*, 2011]. Or, dans les travaux actuels, les deux étapes de construction des objets (segmentation) puis de classification sont toujours faites de manière indépendante. De ce fait, si un objet a été mal construit/extrait lors de l'étape

de segmentation, il ne pourra pas être classé correctement, car il ne correspondra pas (spatialement et spectralement) aux caractéristiques attendues pour ce type d'objet.

Afin de pallier à ce problème, une interaction et une collaboration entre les deux étapes de segmentation et de classification deviennent nécessaires tout en se basant sur les connaissances expertes du domaine. Pour ce faire, nous nous sommes focalisés sur l'étape de segmentation qui représente la pierre angulaire de l'analyse d'image objets. Ceci est dû à la dépendance de la classification à cette étape de segmentation qui permet de construire les objets à classer. Les approches que nous proposons dans cette thèse se basent donc sur les approches de segmentation qui favorisent le plus la notion de collaboration, à savoir, une approche basée sur la segmentation par croissance de région qui offre un environnement propice à la collaboration et à l'intégration de connaissances vue son architecture itérative et hiérarchique (i.e., chapitre 2 et 3) et une approche basée sur la segmentation par collaboration contour région qui est une approche de nature collaborative, i.e., chapitre 4 et 5.

1.6 Conclusion

Dans ce chapitre, nous avons présenté les différentes approches d'analyse d'images. Les approches classiques, bien qu'elles offrent de très bons résultats pour les images à basses et moyennes résolutions, présentent des limites lors du passage vers les images à hautes et très hautes résolutions. Les approches basées objets ont été introduites pour palier à ces limites. Bien qu'elles présentent des résultats intéressants, elles restent tributaires de la phase de construction des objets. C'est dans ce cadre que nous proposons de nouvelles approches collaboratives qui feront l'objet des prochains chapitres.

Croissance de région hiérarchique sémantique : une approche collaborative pour la segmentation et la classification d'objets.

2.1 Introduction

L'analyse d'images basée objets est composée essentiellement de deux étapes. La première consiste en la construction des objets qui est généralement une segmentation. La deuxième étape permet l'identification des objets extraits lors de la première étape. L'analyse d'images basée objets se base sur les régions extraites lors de l'étape de segmentation mais aucune interactivité n'existe entre cette étape et celle de l'identification. Or une région mal segmentée est généralement mal identifiée vu que les caractéristiques extraites de cette région peuvent être erronées. En effet, le passage du pixel vers l'objet permet l'utilisation des informations haut niveau donnant un aspect sémantique à l'objet telles que la forme des régions et la surface. Ces informations seront obligatoirement erronées sur un objet mal segmenté et déformé ce qui génère

une identification erronée. Il est donc utile de permettre une collaboration entre la segmentation et la classification.

Une des approches de segmentation qui favorise le plus la notion de collaboration est l'approche région. En effet, son architecture itérative permet d'introduire de nouvelles connaissances au fur et à mesure du processus de segmentation. C'est dans ce cadre que nous proposons une approche collaborative entre les algorithmes de croissance de régions et une analyse d'images basée objets. En effet, les algorithmes de croissance de régions présentent quelques limites dont principalement le choix des régions de départ (i.e., germes) ainsi que le critère de croissance. Cette problématique peut être résolue en intégrant des informations provenant d'une classification qui peut renseigner sur l'appartenance des objets tout en évaluant la confiance que nous avons sur cette information. Ceci permettra de donner un aspect sémantique à la segmentation ainsi qu'une hiérarchisation du choix des germes et de la croissance selon un certain critère de confiance.

Dans ce chapitre, nous commençons par présenter les algorithmes de croissance de régions. Nous présentons ensuite notre approche de croissance de régions hiérarchique basée sur la classification.

2.2 Croissance de régions : État de l'art et positionnement

Les algorithmes de croissance de régions sont des algorithmes rapides et robustes, qui permettent dans la plupart des cas d'avoir des résultats satisfaisants [Fan *et al.*, 2005]. Ces algorithmes sont constitués de deux principales phases. La première est l'extraction des germes qui représentent le point de départ de la deuxième phase qui consiste en une croissance des régions permettant la construction des objets.

Lors de ces deux phases, trois paramètres ont une influence primordiale sur le résultat. Le premier concerne la phase de détection des germes de départ et les deux autres concernent la phase de croissance :

- La sélection des germes de départ influe énormément sur le résultat final de la segmentation. En effet, le point ou la région germe doit faire partie de la région ciblée, au risque d'obtenir une segmentation partiellement, voire totalement er-

ronée. En effet, une sur-segmentation peut être le résultat d'une surestimation du nombre de régions. Une sous-estimation peut engendrer une sous-segmentation et même une omission de certaines régions. La sélection des germes peut être automatique ou faite par l'utilisateur.

- L'ordre de traitement des germes de départ et des régions candidates à la croissance est lui aussi très influant sur le résultat final. Ceci a été montré par Mehnert et Jackway [Mehnert et Jackway, 1997]
- Le critère de croissance qui permet de décider s'il y a une croissance ou pas à partir de chaque germe, plus explicitement c'est le critère qui permet de décider si nous devons fusionner le germe avec la région traitée.

Plusieurs approches ont été proposées dans la littérature afin d'améliorer la croissance de régions en proposant plusieurs solutions aux étapes précédemment citées. Nous proposons dans ce qui suit une description de quelques unes de ces approches que nous classons en deux catégories : les approches bas-niveau et les approches sémantiques.

2.2.1 Croissance de régions bas niveau

Dans cette section, nous présentons une panoplie d'approche de croissance de régions bas niveau qui se base sur les critères bas niveau tels que la radiométrie et la texture.

Fan *et al.* [Fan *et al.*, 2001a] proposent une approche automatique de croissance de régions. L'approche proposée combine les résultats d'un algorithme d'extraction de contour basée sur les couleurs et d'un algorithme de croissance de régions. La première étape est l'extraction d'une carte de contours en combinant un détecteur de contour isotrope avec un algorithme de seuillage isotrope. Ensuite, les auteurs proposent d'extraire les germes de départ en se basant sur cette carte. En effet, ils se fondent sur l'hypothèse que chaque contour extrait est représentatif d'un objet de l'image. Les germes seront donc les centroïdes des régions contours. Ces germes seront ensuite recalculés chaque fois que la région croît. L'ordre de traitement des germes de croissance est aléatoire. Pour l'étape de croissance, les auteurs décomposent les pixels formant l'image en deux ensembles : les pixels alloués qui représentent les pixels affectés à une région et fusionnés avec un des germes, et les pixels non alloués qui sont les pixels non encore affectés à aucun des germes. Ces pixels sont fusionnés avec le germe qui minimise une distance basée sur les couleurs.

Shih et Cheng [Shih et Cheng, 2005] quant à eux présentent une approche automatique pour l'extraction des germes de départ. L'algorithme de croissance de régions qu'ils proposent permet la segmentation des images couleur. L'approche est composée de quatre étapes. La première étape est la transformation de l'image RGB en entrée vers l'espace de couleurs YC_bC_r . La deuxième étape est le choix automatique des germes. Pour ce faire, les auteurs se basent sur trois critères qui doivent être satisfaits par un pixel pour qu'il puisse être considéré comme germe. D'abord, le pixel doit être un maximum local. Ensuite, pour une région donnée, au moins un germe doit être produit afin d'extraire cette région. Cette condition a été traduite en utilisant la distance euclidienne relative qui doit être maximale pour le pixel choisi pour être germe par rapport à ces huit voisins et elle doit être minimale par rapport à un seuil donné. Enfin, les germes pour les différentes régions doivent être déconnectés. La troisième étape est l'application de la croissance de régions à partir des germes extraits dans l'étape 2. Pour gérer l'ordre de traitement des pixels candidats à la croissance, les auteurs regroupent tous les pixels non traités dans une liste et la trient par rapport à une distance d fondée sur la différence des couleurs dans l'espace YC_bC_r des germes et des régions adjacentes à ces germes. La dernière étape s'attaque au problème de la sur-segmentation en procédant à une étape de fusion des régions voisines et dont la distance ne dépasse pas un seuil choisi par l'utilisateur. Cette approche reste dépendante du choix des seuils précédemment mentionnés. D'autre part, les deux derniers critères sont rarement satisfaits sans connaissances a priori sur les régions existantes dans l'image.

Cui *et al.* [Cui *et al.*, 2008] introduisent une solution aux étapes d'extraction de germes et aux choix des propriétés de croissance afin d'améliorer la croissance de régions. Pour la première étape d'extraction des germes, ils proposent d'utiliser le détecteur de coin de Harris pour détecter les germes de départ. Ce détecteur est invariant aux rotations, aux translations, aux changements d'échelles et au bruit dans l'image. Il est basé sur une fonction d'auto corrélation du signal qui calcule les changements du signal dans les différentes directions. Cependant, les résultats de l'extraction restent tributaires des paramètres de l'algorithme de Harris. Pour la phase de gestion des critères de croissance, les auteurs mettent l'accent sur l'importance des seuils utilisés afin de décider si le pixel traité doit être fusionné avec le germe ou pas. Pour automatiser ce seuil, ils utilisent les modèles de nuages. Pour chaque germe, ils extraient une fenêtre de taille $w * w$ pixels centrée sur ce dernier et y extraient les concepts

qualitatifs permettant de décider si un pixel peut être fusionné avec le germe ou non.

Erickson [Ericksson, 2004] propose un algorithme de croissance de régions par mouvement brownien pour l'extraction des couronnes des arbres. Cet algorithme est composé de quatre étapes. La première étape est une étape de prétraitement qui commence par un seuillage permettant de se débarrasser de tout ce qui diffère des arbres (par exemple le sol nu, les roches). À partir de l'image seuillée, la matrice de distance D qui représente la distance de chaque point de l'image par rapport à l'arrière plan est créée. L'étape de prétraitement se termine par le lissage de la matrice D par un filtre gaussien. La deuxième étape est l'extraction des germes. Elle commence par l'utilisation d'un masque pour la détection des maximums locaux. L'hypothèse posée par Ericksson est que les centres des arbres sont les points les plus lumineux. Les maximums locaux représentent donc les sommets des arbres et les centres des couronnes. La troisième étape est la croissance des couronnes d'arbres. Cette étape est entamée par la création de l'image *Numpass* avec la taille de l'image initiale. Cette image est créée en partant de l'ensemble des maximums locaux extraits. Pour chacun d'entre eux, le phénomène de mouvement brownien qui représente le mouvement aléatoire d'une particule en suspension dans un fluide [Eriksson, 2005, Ericksson, 2004] est simulé. En effet, il s'agit de sommer les $2D$ vecteurs aléatoires, la somme des n vecteurs aléatoires est la position de la particule après n étapes. Pour s'assurer que la particule ne dépasse pas les limites des régions et atterrit dans l'arrière plan et avant d'ajouter le vecteur à la somme, il faut vérifier que l'ajout du vecteur ne propulse pas la particule vers un pixel d'arrière plan i.e., un pixel noir. L'image *Numpass* présente alors le nombre de passages de la particule pour chaque pixel. Finalement la quatrième étape est la limitation des frontières entre les couronnes qui chevauchent.

Cet algorithme dépend de plusieurs paramètres dont la taille du masque permettant l'extraction des germes et dont un mauvais choix peut engendrer une sous ou une sur segmentation des couronnes d'arbres. Le nombre de pas utilisé lors de la croissance par mouvement brownien peut aussi influencer le résultat final.

Jiang *et al.* [Jiang *et al.*, 2011] proposent un nouvel algorithme de croissance de régions basée sur la différence entre les fenêtres de coupes extraites entre deux trames consécutives de séquences d'images afin d'extraire les contours des artères dans les images ultrason. En premier lieu, une fenêtre de différence de coupe est construite en calculant la différence entre deux coupes consécutives. L'étape 2 est l'extraction

des germes de départ. Partant de l'hypothèse que les microbulles dans la lumière se déplacent beaucoup plus rapidement que les artères et les tissus qui les entourent, les auteurs supposent qu'entre deux trames consécutives de séquences d'images échographiques qui ont un taux de rafraîchissement élevé, les tissus et les artères restent figées par rapport au sang qui est en mouvement. L'extraction des germes de départ se fait alors en détectant les maximums locaux de l'image représentant le déplacement du sang calculé dans la première étape. La troisième étape est l'étape de croissance multi-seuil qui combiné avec la différence entre les fenêtres de coupes et les intensités des voisinages du germe essaie de réduire le manque de robustesse des approches de croissance standard par rapport aux taches dans l'image.

2.2.2 Croissance de régions sémantique

Dans cette section, nous présentons les travaux qui ont intégré la sémantique dans les algorithmes de croissance de régions. Ces approches proposent d'intégrer des connaissances pour guider le processus de segmentation.

Fan *et al.* [Fan *et al.*, 2001b] proposent un algorithme automatique sémantique d'extraction d'objets basé modèle et ils proposent comme application la détection des personnes. Comme l'illustre la figure 2.1, l'algorithme est composé de trois étapes. La première étape est la création d'un ensemble de régions homogènes en intégrant les résultats obtenus par une croissance de régions et une détection de contour. La deuxième étape est l'extraction des objets germes à partir des objets extraits de la première étape. Pour ce faire, ils utilisent une carte de couleur de peau afin d'extraire les objets visages comme germe de départ. La troisième étape est l'agrégation des objets connectés à l'ensemble des germes pour générer les objets sémantiques correspondant à un ensemble de modèles.

Un des inconvénients de cette approche est que l'extraction des germes est dépendante des détecteurs d'objets et qu'elle n'affecte pas de degré de confiance aux germes en question. Un autre inconvénient consiste en la sur-détection de germes qui peut affecter la phase de croissance.

Yu et Clausi [Yu et Clausi, 2008] proposent une méthode de segmentation d'image nommée itérative croissance de régions sémantique (IRGS), qui se caractérise par deux aspects. Tout d'abord, ils utilisent des fonctions de pénalité de contour pour approximer

le modèle de contexte spatial des champs de Markov aléatoire (MRF). Deuxièmement, ils utilisent une technique de croissance de régions à la recherche des solutions à ces fonctions objectives. En outre, les IRGS donnent la possibilité de construire une représentation hiérarchique du contenu des images et permettent d'intégrer des connaissances du domaine dans le processus de segmentation. La problématique majeure dans cette approche réside dans le choix des germes de départ qui se base sur la détection des maxima locaux. La phase de croissance est assimilée à une hiérarchie de clustering. Le premier niveau est une sur-segmentation. Ensuite, une itération sur ces régions essaie de minimiser le nombre de cluster. Chaque paire de cluster est examinée, si l'énergie de la fusion minimise l'énergie précédente alors la fusion est justifiée, sinon la fusion est annulée. L'ordre de fusion est définie en fusionnant d'abord les paires qui minimisent le plus l'énergie. Les connaissances du domaine sont intégrées dans la définition de l'énergie.

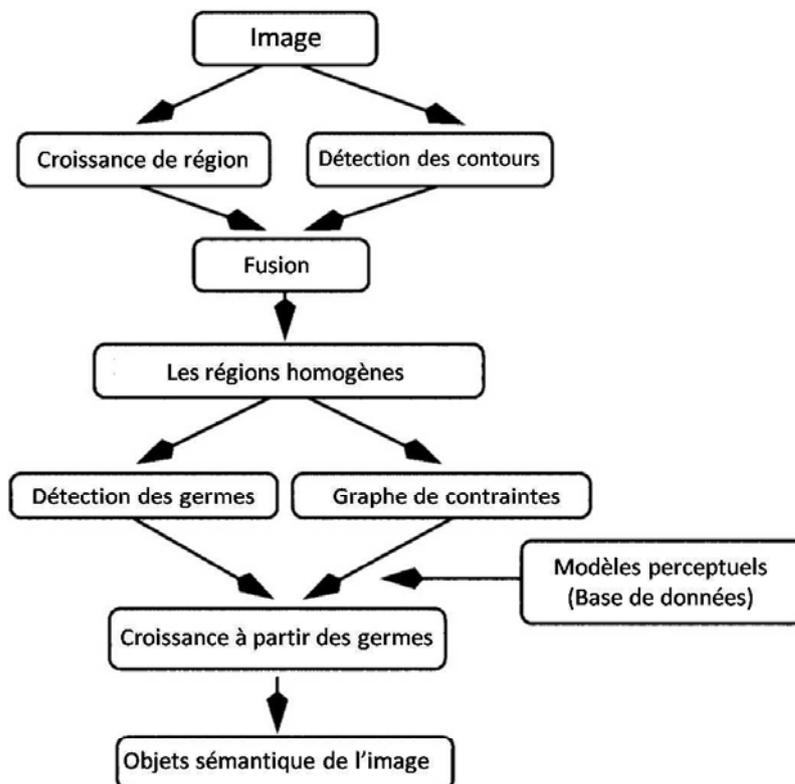


Figure 2.1 — Architecture de l'approche sémantique de Fan *et al.*
 [Fan *et al.*, 2001b].

Athanasiadis *et al.* [Athanasiadis *et al.*, 2007] proposent de réduire le fossé entre les connaissances sémantiques et les connaissances bas niveau en utilisant un algorithme de croissance de régions sémantique pour l'analyse de documents multimedia. Les auteurs commencent par l'extraction des germes en utilisant une classification floue mixée avec un ensemble de règles et de contraintes. En effet, la sélection des germes se fait sur deux critères : (i) Le germe doit avoir un score d'appartenance flou maximal $h(La)$ supérieur à un seuil T_{seed} . (ii) $h(La)$ doit être supérieur à la somme des scores d'appartenance aux autres classes. L'inconvénient majeur de cette approche réside dans le choix du seuil T_{seed} qui est variable et dépend de l'image traitée. Les auteurs cherchent à estimer ce seuil en évaluant le nombre de seuils candidats mais ce choix reste estimatif. En ce qui concerne la croissance proposée, elle utilise une connaissance basée sur la classification floue de l'image et qui informe sur la classe des différents germes. Cependant, la croissance se focalise spécialement sur les résultats de la classification et n'intègre pas vraiment de connaissances expertes. La figure 2.2 illustre le processus de l'approche proposée.

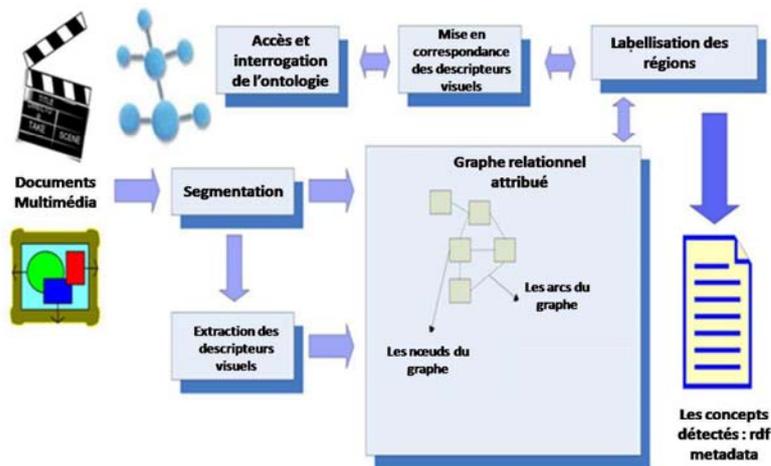


Figure 2.2 — Architecture de la croissance sémantique d'Athanasiadis *et al.*
 [Athanasiadis et Kollias, 2007].

2.2.3 Synthèse et étude comparative

Suite à l'étude des différentes méthodes de croissance, que ce soit à bas niveau ou sémantique, nous notons que la plupart des approches mentionnées se sont concentrées sur la recherche d'une solution aux trois étapes précédemment mentionnées, i.e. le choix des germes de départ, l'ordre de traitement de ces germes et le critère de croissance. Le tableau 6.3 résume l'ensemble des solutions proposées par les travaux cités pour ces trois étapes.

Des hypothèses sont souvent utilisées comme critère de choix de ces germes (par exemple les maxima locaux comme point centraux pour les objets d'intérêts, la détection des coins par des détecteurs de coins tel que Harris). Le problème de ces solutions reste dans les hypothèses elles mêmes qui manque généralement de robustesse dans le sens où elles ne sont pas toujours vraies et qu'elles sont très sensibles au bruit dans l'image, ce qui risque de générer une surestimation du nombre de germes. D'autres approches se basent sur une première étape de décomposition de l'image en un ensemble d'objets dont les centroïdes seront identifiés comme l'ensemble des germes. Le problème de ces approches réside dans la dépendance à la phase de décomposition qui peut induire à une surestimation ou une sous-estimation du nombre des germes. Une troisième solution proposée est l'utilisation d'une classification comme étape initiale afin d'extraire les objets les plus pertinents comme germes. Le problème avec cette solution proposée par Athanasiadis [Athanasiadis *et al.*, 2007] est qu'elle est dépendante du seuil qui permet la sélection des germes à chaque itération.

Concernant l'ordre de traitement des germes, deux solutions sont les plus utilisées. La première solution est un choix aléatoire qui fait que le résultat final de la segmentation change à chaque fois. La deuxième solution consiste à trier l'ensemble de germe selon sa distance locale avec l'ensemble de ces voisins. Cette solution favorise donc les germes qui sont les plus proches de leurs voisins mais elle reste sensible au bruit. Pour le choix du critère de croissance, la majorité des approches utilise les connaissances bas niveaux et généralement la radiométrie comme base de la croissance. Concernant les approches de croissance sémantique proposées, certaines ne permettent pas une flexibilité pour l'intégration des connaissances [Yu et Clausi, 2008] alors que d'autres restent spécifiques à des objets particuliers.

Pour résumer, le choix des germes continue de freiner les solutions proposées pour la croissance de régions. La gestion de l'ordre de traitement des germes est aussi un

Tableau 2.1 — Étude comparative des méthodes de croissance de régions

Les approches	Extraction des germes	Ordre de croissance à partir des germes	Croissance
Fan et al. [Fan <i>et al.</i> , 2001a]	Centroïdes des contours	Aléatoire	Radiométrie
Shih and Cheng [Shih et Cheng, 2005]	Radiométrie + règles	Trie les germes par distance entre le germe et ces voisins	Radiométrie
Cui et al. [Cui <i>et al.</i> , 2008]	Détecteur de coin de Harris	aléatoire	Radiométrie + modèle des nuages
Erickson [Ericksson, 2004],	Maxima locaux	Ordonnée aléatoirement et candidats à la croissance ordonnés par distance par rapport aux germes	Radiométrie
Jiang <i>et al.</i> [Jiang <i>et al.</i> , 2011]	Maxima locaux	Croissance aléatoire multi-seuil	Radiométrie
Fan <i>et al.</i> [Fan <i>et al.</i> , 2001b]	Centroïdes des objets visages détectés à partir d'une première étape de segmentation coopérative région contour	Aléatoire	Règles définissant un modèle
Yu et Clausi [Yu et Clausi, 2008]	Maximums locaux	Les germes dont la fusion minimise une énergie	Sémantique en intégrant les connaissances dans l'énergie utilisée lors de la fusion,
Athanasiadis [Athanasiadis et al., 2007]	Classification floue	Aléatoire sur chaque itération	Se base sur la classification flou

problème qu'il faut gérer vu son importance sur la solution finale. Enfin, le manque de connaissances fondées sur la nature des objets traités dans le processus de croissance est aussi un inconvénient majeur de ces approches. Même si les approches sémantiques proposées ont essayé de palier à cet inconvénient, l'intégration de la connaissance reste restreinte et n'est pas généralisée. Nous proposons dans la suite de ce chapitre une solution à ces trois étapes.

2.3 Principe de l'approche proposée

Comme nous l'avons présenté dans le premier chapitre, le problème majeur des approches d'analyse d'image est le manque d'interaction entre l'étape de segmentation et celle de classification. En analysant les différents travaux présentés dans la littérature, nous avons conclu qu'une interaction entre ces deux étapes serait une solution à ce problème.

La première approche que nous proposons dans ce chapitre se base sur la segmentation par croissance de région. Cette technique de segmentation présente une architecture itérative et propice à la collaboration, nous avons aussi constaté que les trois étapes primordiales la composant souffrent du manque de l'intégration de la sémantique et des connaissances expertes pouvant permettre de remédier aux lacunes de cette dernière. Or les approches d'analyse d'images basées objets facilitent l'intégration de nouveaux attributs et de connaissances sémantique facilitant la description des objets. Ce qui nous a amené à constater que la coopération entre un algorithme de croissance de régions et une approche de classification au sein d'une approche d'analyse d'images basées objets pourrait présenter une solution aux inconvénients des deux approches. Dans ce cadre, nous proposons une approche collaborative hiérarchique pour la croissance de régions basée sur la classification. Cette approche introduit une solution au problème de choix des germes de départ ainsi qu'au problème du choix de l'ordre de traitement de ces derniers. Elle permet aussi l'intégration sémantique dans le processus de croissance. Cette approche est décomposée en deux étapes :

- Une première étape de prétraitement qui représente une phase de préparation des données de départ pour la phase suivante. Elle consiste à décomposer l'image en un ensemble d'objets homogènes selon des critères bas niveaux tels que la radiométrie et la texture. Une fois les objets homogènes extraits, nous classons

l'image à travers le calcul d'un score, appelé score de similarité, pour chaque objet, en se basant sur des connaissances bas niveaux fournies par l'expert. Ce score permettra d'introduire la notion de confiance qui renseigne sur le degré de validité de l'étiquetage de chaque objet (i.e. sur l'appartenance de l'objet à chaque classe) et permettra donc de créer une hiérarchie de croissance basée sur cette confiance.

- Une seconde étape, correspondant à la croissance hiérarchique, est une étape itérative où chaque itération est composée de quatre phases.
 - La première phase est la sélection de l'ensemble des germes en se basant sur les scores de similarité. En effet, cet ensemble représente les objets non encore traités et qui maximisent le score de similarité.
 - La deuxième phase consiste à effectuer, pour chaque objet, un ensemble de traitements en se basant sur des contraintes spatiales, si elles existent, afin d'élaguer cet ensemble et de remettre en cause la classification initiale.
 - La troisième phase est l'extraction de la zone de croissance selon la classe du germe.
 - La dernière phase est la phase de croissance.

La figure 2.3 illustre l'architecture globale de l'approche que nous détaillons dans ce qui suit.

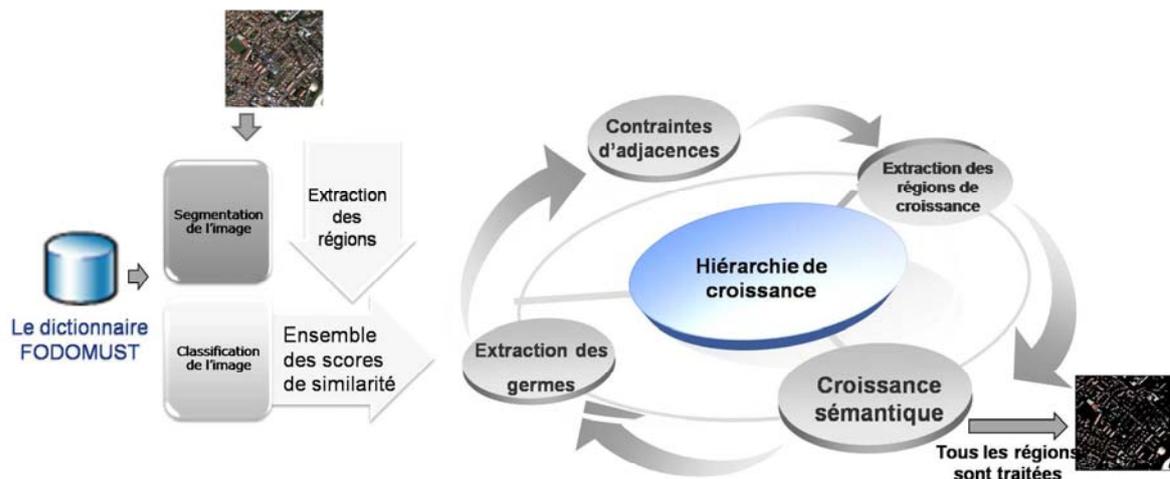


Figure 2.3 — Workflow de l'approche.

2.3.1 Prétraitement

2.3.1.1 Segmentation

Le choix de l'algorithme de segmentation n'est pas très important dans cette approche tant qu'il vérifie le critère de sur-segmentation. En effet, vues les propriétés des algorithmes de croissance de régions qui se basent sur la fusion des fragments d'un objet afin de détecter l'objet tout entier, il est évident qu'une sous-segmentation de l'image implique une perte de certains objets. Le choix de la sur-segmentation s'impose. Nous avons choisi l'algorithme de ligne de partage des eaux (i.e., Watershed) qui permet une sur-segmentation de l'image, permettant de déterminer un ensemble de régions de départ.

2.3.1.2 Classification

Notations - Soit \mathcal{R} l'ensemble des régions r_i obtenu à partir de la segmentation. $\mathcal{R} = \{r_i\}_{i \in [1, N_R]}$ où N_R représente la cardinalité de \mathcal{R} . Soit \mathcal{C} l'ensemble des classes présentes dans l'image. $\mathcal{C} = \{c_j\}_{j \in [1, N_C]}$ où N_C représente la cardinalité de \mathcal{C} . Soit \mathcal{A} l'ensemble des attributs identifiant une classe. $\mathcal{A} = \{a_k\}_{k \in [1, N_k]}$ où N_k représente la cardinalité de \mathcal{A} .

Nous proposons ici une classification permettant d'affecter à chacune des régions de \mathcal{R} une classe c_j en calculant un score de confiance d'appartenance de \mathcal{R} à c_j , par rapport à l'ensemble des classes présentes dans l'image. Ce score permettra d'évaluer la légitimité des régions en se basant sur les connaissances fournies par l'expert. Nous utilisons le score de similarité proposé par [Derivaux *et al.*, 2007]. Il est basé sur une approche orientée attribut car il utilise les connaissances bas niveaux sur l'image, formalisées sous forme de descripteurs bas niveaux. Ce score permet de vérifier la validité des valeurs des attributs d'une région selon les intervalles définis par l'expert dans la base de connaissances. La mesure de similarité locale compare les valeurs des attributs d'une région avec les attributs de l'objet à classer.

Définition 1. (Degré de validité) : Le degré de validité entre une classe c_j et une région r_i pour un attribut a_k , que l'on note $Valid(a_k, c_j, r_i)$, est défini comme suit :

$$Valid(a_k, c_j, r_i) \begin{cases} 1 & \text{si } v(r_i, a_k) \in [\min(c_j, a_k), \max(c_j, a_k)] \\ \frac{v(r_i, a_k)}{\min(c_j, a_k)} & \text{si } v(r_i, a_k) < \min(c_j, a_k) \\ \frac{\max(c_j, a_k)}{v(r_i, a_k)} & \text{si } v(r_i, a_k) > \max(c_j, a_k) \end{cases} \quad (2.1)$$

où $v(r_i, a_k)$ est la valeur de l'attribut a_k pour la région r_i . Ce degré de validité évalue la correspondance entre une région r_i et une classe c_j pour un attribut a_k . En effet, si la valeur de l'attribut a_k pour la région r_i appartient à l'intervalle spécifié pour la classe c_j , alors, ce degré est maximal, i.e., égal à 1. Sinon, plus on s'éloigne de cet intervalle, plus le degré de validité diminue et tend vers 0.

Définition 2. (Score de similarité) : Le score de similarité, noté $Sim(r_i, c_j)$, est calculé en fonction de la validité entre la région r_i et la classe c_j de chaque attribut a_k . Le calcul est pondéré par un poids $w(k, c_j)$, qui permet de gérer l'importance de chaque attribut pour les différentes classes. Ce score représente la validité globale d'une région par rapport à une classe en se basant sur tous les attributs. La formule du calcul du score de similarité est donc la suivante :

$$Sim(r_i, c_j) = \frac{\sum_{a \in A} w(a_k, c_j) Valid(a_k, c_j, r_i)}{\sum_{a \in A} w(a_k, c_j)} \quad (2.2)$$

Définition 3. (Ensemble de similarité) : Nous définissons l'ensemble SIM comme étant l'ensemble des scores de similarité de toute région $r_i \in \mathcal{R}$ par rapport à toute classe $c_j \in \mathcal{C}$.

$$SIM = \{Sim(r_i, c_j) | r_i \in \mathcal{R} \text{ et } c_j \in \mathcal{C}\} \quad (2.3)$$

2.3.2 Croissance hiérarchique sémantique

Après la phase de prétraitement décrite ci-dessus, nous procédons à l'étape de classification hiérarchique sémantique qui consiste en un traitement itératif permettant à partir des ensembles de régions \mathcal{R} , des classes \mathcal{C} et des scores de similarité SIM , de créer une hiérarchie de croissance basée sur la confiance en chaque région. La création de la hiérarchie est précédée par un calcul basé sur les scores de similarité que nous explicitons dans ce qui suit.

Définition 4. Pour une région $r_i \in \mathcal{R}$, nous définissons l'ensemble des classes qui maximisent le score de similarité $Sim(r_i, c)$ parmi toutes les classes $c \in \mathcal{C}$. Nous notons

$\delta(r_i)$ cet ensemble :

$$\delta(r_i) = \arg \max_{c \in \mathcal{C}} \text{Sim}(r_i, c) \quad (2.4)$$

Définition 5. Pour chaque région $r_i \in \mathcal{R}$, nous définissons $S_{max}(r_i)$ et $C_{max}(r_i)$ comme suit :

$$C_{max}(r_i) = \begin{cases} \text{random}(\delta(r_i)) & \text{if } |\delta(r_i)| > 1 \\ \delta(r_i) & \text{sinon} \end{cases} \quad (2.5)$$

$$S_{max}(r_i) = \begin{cases} \text{Sim}(r_i, C_{max}(r_i)) & \text{if } |\delta(r_i)| = 1 \\ 0 & \text{sinon} \end{cases} \quad (2.6)$$

$S_{max}(r_i)$ représente le score de similarité maximal de la région r_i pour l'ensemble des classes de \mathcal{C} . Dans le cas où $\delta(r_i)$ comporte plus qu'une classe, nous déduisons qu'il y a une confusion et que cette région n'est plus une région de confiance mais une région conflictuelle. Dans ce cas, $C_{max}(r_i)$ prendra arbitrairement l'une des classes de $\delta(r_i)$ et $S_{max}(r_i)$ aura la valeur 0. Dans le cas où l'ensemble contient une valeur unique, alors cette dernière sera affectée à $C_{max}(r_i)$ et $S_{max}(r_i)$ sera le score de similarité $\text{Sim}(r_i, C_{max}(r_i))$ de la classe C_{max} pour la région r_i et cette région aura comme classe $C_{max}(r_i)$.

Le calcul des $C_{max}(r_i)$ et $S_{max}(r_i)$ servira de départ à l'algorithme itératif de la classification hiérarchique sémantique que nous proposons. Le diagramme de la figure 2.4 illustre les étapes de cet algorithme. En effet, chaque itération de cet algorithme concerne un niveau de croissance de la hiérarchie. Chacune de ces itérations est composée de deux phases principales (les phases 1 et 2 du diagramme), à savoir l'extraction des germes et la croissance sémantique. Dans la première phase, nous commençons par extraire les régions de départ que nous appelons germes. Ces derniers représentent les régions de confiance parmi l'ensemble des régions candidates. Ensuite dans la deuxième phase, en se fondant sur des connaissances expertes, et selon la classe d'appartenance du germe, nous procédons à une croissance sémantique à partir des germes déterminés dans la première phase.

Pour une itération k ($k \geq 1$), nous désignons par $RegionsCandidates_{k-1}$ l'ensemble des régions candidates à l'extraction des germes, par $RegionsContraintes_{k-1}$ l'ensemble des germes déjà traités, par $RegionsGermes_k$ l'ensemble des germes extraits et par $RegionsFusionnees_k$ l'ensemble des régions fusionnées.

Les phases 3 et 4 permettent de mettre à jour les ensembles $RegionsCandidates_k$

et $RegionsContraintes_k$. Nous détaillons dans ce qui suit les deux phases constituant une itération.

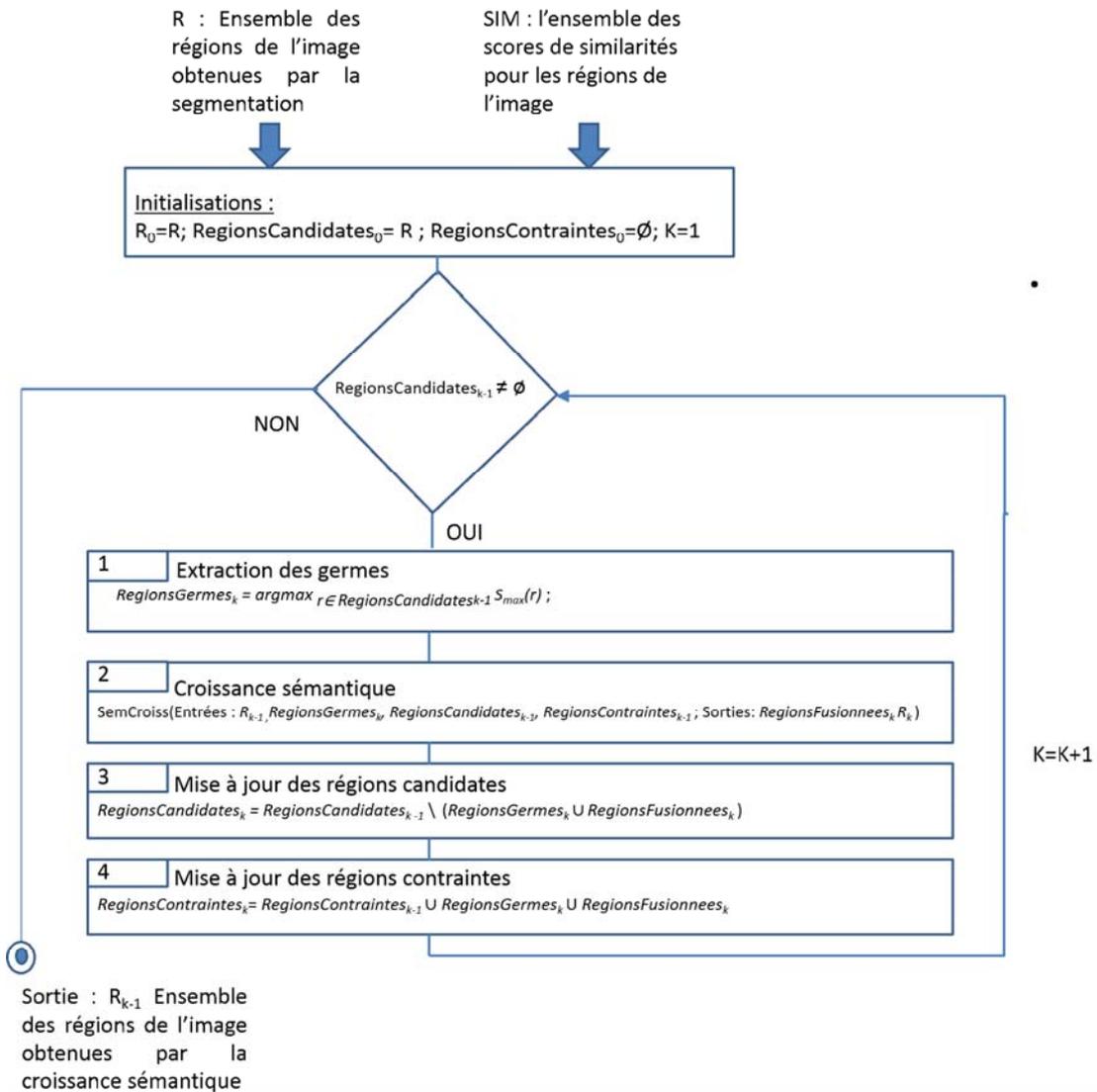


Figure 2.4 — Diagramme de croissance de régions sémantique.

2.3.2.1 Extraction des germes

Le choix des germes de départ est très important pour le processus de la croissance sémantique. Visant à exploiter toutes les informations disponibles et souhai-

tant converger vers une approche sémantique, le choix des germes reposera sur la confiance que nous avons en les régions non encore traitées dans l'image. L'ensemble des germes de niveau k ($RegionsGerme_k$) sera alors celui qui maximise la confiance et plus précisément le score de similarité. Ces régions seront extraites à partir de l'ensemble $RegionCandidates_{k-1}$ des régions non encore traitées dans les niveaux précédents de la hiérarchie. Les germes extraits à ce niveau sont les régions r_i qui maximisent $S_{max}(r_i)$. Nous notons formellement :

Définition 6. (*Germes*) :

$$RegionsGerme_k = \arg \max_{r \in RegionsCandidates_{k-1}} S_{max}(r). \quad (2.7)$$

La figure 2.5 illustre sur un exemple artificiel le processus d'extraction de germe pour la première itération. La figure présente l'ensemble des objets composant l'image. Les deux objets A et K ayant le score de similarité maximal S_{max} égal à 0.95 seront le point de départ de la première étape de croissance. Chaque objet aura une étape de croissance spécifique à sa classe d'appartenance C_{max} . Pour le premier niveau de croissance, aucune contrainte n'est admise.

2.3.2.2 Intégration des connaissances spatiales d'adjacence

Les attributs contextuels permettent de caractériser les informations spatiales entre les objets qui représentent des informations très intéressantes vue leur stabilité [Bloch, 2002, Inglada et Michel, 2009]. Les relations d'adjacence sont parmi les informations spatiales les plus utilisées. Dans ce cadre, nous présentons un algorithme permettant d'intégrer ces informations dans le processus de croissance sémantique. Pour ce faire, nous modélisons les relations d'adjacence sous forme de contraintes informant sur la possibilité de cohabitation entre les classes. Pour chaque classe, nous introduisons une liste de classes contraintes qui représente l'ensemble des classes qui ne peuvent pas être adjacentes avec la classe traitée. Dans chaque niveau de la hiérarchie de croissance, cet ensemble de contraintes permettra de réévaluer la confiance en chaque germe en se fondant sur ces contraintes. Ces contraintes permettront aussi de réévaluer la classification initiale obtenue lors de la phase de prétraitement. Pour modéliser



Figure 2.5 — Illustration de l'extraction des germes sur une image artificielle.

l'intégration de ces connaissances nous définissons les ensembles suivants :

Définition 7. (Voisinage) : Soit $Voisinage(r_i)$ l'ensemble des régions r_j adjacentes à r_i .

$$Voisinage(r_i) = \{r_j \in R | r_i \text{ et } r_j \text{ sont adjacentes}\} \quad (2.8)$$

Définition 8. (Contraintes spatiales) : Soit $CS(c_i)$ l'ensemble des classes c_j qui représente les contraintes spatiales de la classe c_i .

$$CS(c_i) = \{c_j \in C | c_i \text{ et } c_j \text{ ne peuvent pas être voisines}\} \quad (2.9)$$

Pour chaque itération de la hiérarchie, nous vérifions la compatibilité du germe avec

son voisinage. Si deux germes voisins du même niveau de la hiérarchie (i.e. ayant le même S_{max}) sont assignés à deux classes qui ne peuvent pas cohabiter, alors il existe une ambiguïté dans la phase de classification et ces deux germes deviennent conflictuels. Dans ce cas, ils perdent leur place dans la hiérarchie et le score d'appartenance de ces germes aux deux classes de conflit est remis à zéro. Les deux régions sont alors retirées de l'ensemble des germes et sont affectées de nouveau à l'ensemble des candidats non encore traités.

Pour les régions candidates voisines, nous vérifions si la classe d'appartenance de cette région appartient à l'ensemble des contraintes. Dans ce cas, nous mettons à zéro le score de similarité d'appartenance de cette région à la classe d'appartenance tant qu'elle appartient à l'ensemble des classes contraintes.

2.3.2.3 Croissance sémantique

Nous présentons dans ce qui suit le principe de l'algorithme de croissance sémantique (la fonction *SemCroiss* appelée à la phase 2 du diagramme 2.4) qui prend comme entrée l'ensemble des germes, des régions candidates à la croissance et des régions contraintes. Il permet la fusion de chaque germe avec ses régions voisines en se fondant sur les connaissances expertes spécifiques à la classe du germe et en prenant en compte les régions contraintes. Dans ce chapitre, nous présentons le formalisme de l'algorithme sans présenter les détails. En effet, les différentes phases de l'algorithme de croissance sont spécifiques à la classe du germe. Sans connaissances a priori sur cette classe, la croissance sémantique n'a plus lieu d'être. La croissance sémantique est illustrée par l'algorithme 2.

Cet algorithme itère sur l'ensemble de sgermes appliquant pour chacun deux fonctions (*ExtraireZoneCroissance* et *Croissance*) mettant en œuvre les deux principaux procédés composant la croissance sémantique, à savoir l'extraction de la zone de croissance dans un premier lieu et la croissance au sein de cette zone dans un second lieu.

La phase de croissance prend en considération pour chaque itération l'ensemble des objets déjà extraits. Ces objets présentent des contraintes pour la croissance des germes dans les itérations suivantes. Ces contraintes sont basées sur le principe de la confiance. En effet, un germe ayant une confiance inférieure et donc un score de similarité inférieur, ne peut pas modifier la classe d'appartenance d'un objet qui lui

Algorithme 1 : Intégration des connaissances spatiales d'adjacence

Données : $RegionsGerme_k$, SIM and $RegionsCandidates_{k-1}$.

Résultats : $RegionsGerme_k$, SIM and $RegionsCandidates_{k-1}$.

```

1  début
2  |   pour chaque  $r_i \in RegionsGerme_k$  faire
3  |   |   pour chaque  $r_j \in (Voisinage(r_i) \cap RegionsCandidates_{k-1})$  faire
4  |   |   |   tant que  $C_{max}(r_j) \in CS(C_{max}(r_i))$  faire
5  |   |   |   |    $Sim(r_j, C_{max}(r_j)) = 0$ ;
6  |   |   |   |   fin
7  |   |   |   fin
8  |   |    $Flag = 0$ ;
9  |   |   pour chaque  $r_j \in (Voisinage(r_i) \cap RegionsGerme_k)$  faire
10 |   |   |   si  $C_{max}(r_j) \in CS(C_{max}(r_i))$  alors
11 |   |   |   |    $Sim(r_j, C_{max}(r_j)) = 0$ ;
12 |   |   |   |    $RegionsGerme_k = \{RegionsGerme_k \setminus \{r_j\}\}$  :
13 |   |   |   |    $RegionsCandidates_{k-1} = RegionsCandidates_{k-1} \cup \{r_j\}$ ;
14 |   |   |   |   si  $S_{max}(r_i) == S_{max}(r_j)$  alors
15 |   |   |   |   |    $Flag = 1$ ;
16 |   |   |   |   fin
17 |   |   |   fin
18 |   |   si  $Flag == 1$  alors
19 |   |   |    $Sim(r_i, C_{max}(r_i)) = 0$ ;
20 |   |   |    $RegionsGerme_k = RegionsGerme_k \setminus \{r_i\}$ ;
21 |   |   |    $RegionsCandidates_{k-1} = RegionsCandidates_{k-1} \cup \{r_i\}$ ;
22 |   |   |   fin
23 |   |   fin
24 fin
    
```

est supérieur dans la hiérarchie. La figure 2.6 illustre l'intégration de connaissances sous forme de contraintes lors de la phase de croissance hiérarchique. Les deux germes extraits sont les objets M et O. Les deux objets A et K sont des objets déjà classés et présentent donc des contraintes pour les deux germes qui ne peuvent pas croître dans cette direction.

Algorithme 2 : SemCroiss

Données : *RegionsGerme*s : l'ensemble des germes, *RegionsCandidates* :
 l'ensemble des candidats pour la croissance et
RegionsContraintes : l'ensemble des régions de contraintes.

Résultats : *RegionsFusionnees* : l'ensemble des régions fusionnées lors de la
 croissance

```

1  début
2  |   RegionsFusionnees =  $\phi$ ;
3  |   pour chaque  $r_i \in \text{RegionsGerme}s$  faire
4  |       |    $ZC_{r_i} = \text{ExtraireZoneCroissance}(r_i, C_{max}(r_i), \text{RegionsCandidates},$ 
5  |       |       |    $\text{RegionsContraintes}, \text{RegionsFusionnees})$ ;
6  |       |       |    $\text{RegionsFusionnees}_{ees\_avec\_r_i} = \text{Croissance}(r_i, C_{max}(r_i), ZC_{r_i})$ ;
7  |       |       |    $\text{RegionsFusionnees} = \text{RegionsFusionnees} \cup \text{Fusionnees}_{avec\_r_i}$ ;
8  |   fin
9  fin
    
```

L'extraction de la zone de croissance : Connaissant la classe du germe, nous proposons de limiter l'espace de croissance en nous basant sur les connaissances expertes. En effet, celles-ci permettent dans plusieurs cas de réduire l'ensemble des régions candidates à la fusion. Les connaissances expertes peuvent être représentées de plusieurs manières, telles que les ontologies [Derivaux *et al.*, 2007], les règles logiques [Suzuki, 2002], etc. Cette phase est tributaire de la classe traitée. En effet, si nous prenons l'exemple de la classe végétation, rares sont les informations qui peuvent être utiles pour limiter l'espace de recherche. N'ayant ni forme géométrique bien définie, ni superficie limitée, la zone de croissance de cette classe ne peut pas vraiment être limitée. Par contre, si nous prenons l'exemple des classes bâtiment et route, elles possèdent toutes les deux des caractéristiques géométriques et des formes bien définies ce qui permet de limiter l'espace de croissance.

L'appel de la fonction *ZoneCroissance* dans une itération de l'algorithme 2 permet de générer, à partir, d'une région germe r donnée et des régions candidates à la croissance, l'ensemble des régions pertinentes qui peuvent fusionner avec le germe r , que l'on notera ZC_r . Cette fonction prend en considération les régions déjà fusionnées que



Figure 2.6 — Illustration de l'intégration des contraintes dans la croissance hiérarchique sur une image artificielle.

ce soit pour d'autres germes de l'itération courante (i.e. *Fusion*) ou dans les itérations précédentes (i.e. *RegionsContraintes*).

L'ensemble ZC_r serait égal à l'ensemble des candidats privé des régions élaguées dans le cas où la classe $C_{max}(r)$ supporte l'intégration de connaissances expertes comme expliqué précédemment. Dans le cas contraire, ZC_r comporterait toutes les régions candidates sauf celle déjà fusionnée dans l'itération courante (i.e. $ZC_r = RegionCandidates \setminus Fusion$).

Croissance : Contrairement à la croissance de régions de base qui utilise les propriétés bas niveau (radiométrie, texture) sans prendre en compte la nature des objets traités, la croissance de régions sémantique est, quant à elle, une croissance spécifique

pour chaque germe selon sa classe d'appartenance. Prenons l'exemple de la classe route qui possède des propriétés qui la différencient des autres classes. En effet, elle présente une surface homogène, et elle est majoritairement composée d'asphalte et ayant des contours linéaires et parallèles et une largeur variant dans un intervalle bien déterminé. À partir de ces connaissances, nous pouvons détecter la zone de croissance de chaque germe en cherchant sa direction et en utilisant les informations sur sa largeur. L'appel de la fonction *Croissance* dans une itération de l'algorithme 2 permet de croître à partir d'un germe r dans la zone de croissance ZC_r . La détermination des régions pertinentes qui peuvent fusionner avec ce germe r dépendra de la classe du germe et des connaissances expertes la concernant. Cette fonction retournerait l'ensemble *Fusionnees_avec_r* des germes fusionnés avec r . L'union de ces ensembles pour chaque itération formerait l'ensemble de retour de l'algorithme SemCroiss.

2.4 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle approche, d'analyse d'images basée objets, collaborative et sémantique. Cette approche essaie de remédier au manque d'interaction entre les deux étapes majeures des approches d'analyse d'image orientés objet à savoir la construction puis l'identification des objets. Notre approche se base sur une hiérarchie de croissance sémantique permettant de détecter des objets selon la confiance calculée pour chaque région de l'image et en se basant sur des connaissances expertes. Cependant, la partie de croissance sémantique reste dépendante de la classe du germe. Dans le chapitre suivant, nous présenterons une étude des deux classes bâtiment et route. Deux approches de croissance sémantique seront présentées afin d'évaluer la robustesse et l'apport de notre solution.

Croissance hiérarchique sémantique pour l'extraction des bâtiments et des routes

3.1 Introduction

Dans ce chapitre, nous proposons une application de notre approche de croissance de région sémantique présentée dans le chapitre 2. Comme nous l'avons mentionné, la croissance sémantique est dépendante de la classe du germe traité et des connaissances qui le caractérisent. Dans ce cadre, nous entamons ce chapitre par la présentation des connaissances que nous avons sur les images à traiter. Ensuite, nous introduisons deux approches de croissance sémantique pour les deux classes les plus traitées pour l'analyse des zones urbaines, à savoir, la classe bâtiment et la classe route.

3.2 Connaissances du domaine

3.2.1 Le fossé sémantique

Les connaissances expertes représentent une source d'information très importante dans la reconnaissance et la classification d'images. Elles peuvent être un facteur important pour la décision entre deux classes ambiguës ou même une source principale pour tous les systèmes d'analyse. Le problème majeur des connaissances expertes est leur représentation. En effet, il existe une grande différence entre la représentation

sémantique de l'image qui est une représentation haut niveau et sa description numérique (les informations extraites de l'image) qui est de bas niveau. Ce fossé est appelé fossé sémantique. Smeulders *et al.* [Smeulders *et al.*, 2000] le définissent comme étant le manque de coïncidence entre l'information que l'on peut extraire des données visuelles et l'interprétation de ces mêmes données pour un utilisateur dans une situation donnée.

Pour l'analyse des images satellitaires, la première source d'information et de connaissance est l'expert géographe. En se basant sur ses connaissances, il peut fournir une description et une analyse des scènes à interpréter. Cependant, plusieurs travaux ont montré la complexité de la modélisation de ces connaissances afin qu'elles soient exploitables [Smeulders *et al.*, 2000]. En effet, l'expert utilise un langage spécifique au domaine de la photo-interprétation. Les critères tels que la couleur, la taille et la texture sont utilisés. Le passage de cette interprétation de l'expert vers des données directement exploitables par la machine reste une tâche compliquée.

L'exemple présenté par Derivaux [Derivaux, 2009] et Forestier [Forestier, 2010] illustre le problème que pose ce fossé sémantique. En effet, si l'expert photo-interprète peut reconnaître une route immédiatement de part sa couleur grise et sa forme (linéaire et continue dans l'espace), il peut lui être difficile d'exprimer directement des valeurs numériques utilisables par un système informatique. Il existe un fossé sémantique entre sa connaissance et les données à traiter qui ne sont pas de même nature. Pour autant, il est possible d'extraire un ensemble de connaissances de l'expert.

Afin d'y remédier, une modélisation de ces connaissances en un dictionnaire des données a été réalisée par le laboratoire Image, Ville et Environnement dans le cadre du projet Fodomust (ACI Masse de données 2004-2007).

3.2.2 Dictionnaire des données

Le recueil des connaissances du domaine par l'expert en photo-interprétation d'images en milieu urbain est résumé dans un dictionnaire de données réalisé par le laboratoire Image, Ville, Environnement. Ce dictionnaire est un document textuel présenté sous la forme de fiches précisant pour chaque classe d'objets (également appelée concept), susceptible d'être identifiée sur une image THR en milieu urbain, les éléments suivants [Derivaux, 2009, Forestier, 2010] :

1. Une identification pour la classe contenant son nom et sa hiérarchie dans les classes.
2. Une description dans le monde réel :
 - Textuelle : *"La classe bâtiment de l'habitat collectif discontinu appelé aussi immeuble appartient à la catégorie de classe élémentaire bâtiment. Il désigne une construction durable importante, destinée à l'habitation collective en appartements ou à des activités secondaires ou tertiaires, à plusieurs étages. On distingue le plus souvent les immeubles en barre et tour. Une barre est un bâtiment de forme rectangulaire (étroit et allongé au sol) de plus de 4 étages. Une tour est un bâtiment de forme carrée, de volume allongé vers le haut, de plus de 6 étages. Les tours et les barres sont organisées en tissus urbain(TU) discontinu de type Grand Ensemble (GE). Les immeubles de ce type de TU se caractérisent par la régularité des directions principales (parallélisme, orthogonalité). En général, un immeuble : est situé dans un îlot physique (domaine privé); a une emprise au sol de plus de 30m² ; est associé à des surfaces artificialisées (route, parking, espaces verts, aire de jeux) organisées en quartiers (généralement relativement récents, construits après-guerre)"*
 - Visuelle : composée d'un ensemble d'images contenant les objets de la classe comme l'illustre la figure 3.1.
3. Description dans l'image :
 - Nature de l'objet : quelle forme (simple ou composition) prend l'objet selon la résolution spatiale utilisée.
 - Définition textuelle relative à son identification dans l'image. Par exemple : *"Un objet de la classe immeuble est représenté graphiquement par un polygone dont la surface correspond à l'emprise au sol du bâtiment."*
 - Principales relations : notamment l'adjacence avec d'autres classes, s'il existe une notion d'alignement entre des objets de cette classe ainsi que la distance entre des objets de cette classe si cela est approprié. Par exemple : *"Adjacence possible de la classe avec les objets de type végétation et de type route. La distance entre les objets immeubles est moyenne à élevée."*
 - Attributs : type de signature spectrale, longueur, largeur, élongation, surface, critères géométriques, type de texture, etc.
4. Commentaires : liens vers d'autres sources de données ou d'exemples possibles.

<i>THR1 – QB MS (2.8 m)</i>			
- Bâtiment de l'habitat collectif discontinu (type GE)			
	bâtiment de l'habitat collectif	bâtiment de l'habitat collectif de type « barre »	bâtiment de l'habitat collectif de type « tour »

Figure 3.1 — Extrait du dictionnaire des données présentant une description visuelle de la classe "bâtiment : habitat collectif".

3.2.3 Formalisme des données

Le dictionnaire, sous sa forme textuelle, ne permet pas d'intégrer les connaissances dans le traitement informatique. Le manque de structuration de ces informations n'y permet pas un accès facile. Derivaux et Forestier [Derivaux, 2009, Forestier, 2010] ont proposé de modéliser ces informations en utilisant le langage XML. Ce langage permettant de structurer des données semi-structurées est très adapté au dictionnaire. La base de connaissances est apparentée à une ontologie et sa modélisation se fait sur deux niveaux.

3.2.3.1 Hiérarchie de classes

La hiérarchie de classes représente les liens existants entre les différentes classes présentes dans l'image. La figure 3.2 illustre le type de relation qu'on peut avoir entre les classes. En effet, la classe "Bâtiment" est une généralisation des deux classes "Pavillon" et la classe "Immeuble" et inversement ces deux classes sont une spécialisation de la classe "Bâtiment". L'exemple montre donc la spécialisation de la classe la plus générale "Objet" en trois sous-classes "Bâtiment", "Route" et "Végétation". La modélisation hiérarchique permet de gérer les niveaux d'analyse. Plus l'analyse est précise plus on descend dans la hiérarchie afin de collecter plus de détails.

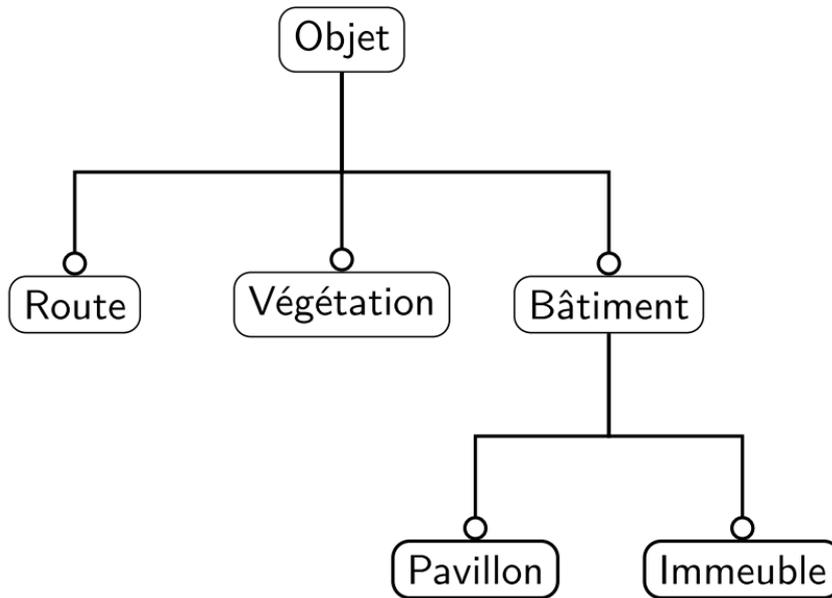


Figure 3.2 — Extrait de la hiérarchie de classes utilisée dans [Forestier, 2010].

3.2.3.2 Caractérisation des classes

Après la définition de la hiérarchie, les auteurs ont décrit l'ensemble des classes par les connaissances issues du dictionnaire de données. Plusieurs types de connaissances peuvent être introduits, à savoir :

- Intervalle sur des attributs quantitatifs. Pour la classe "Bâtiment" par exemple, l'attribut aire va appartenir à un intervalle $[aire_{min}, aire_{max}]$ qui informe sur l'intervalle d'aire que peut avoir un objet de cette classe.
- Méta-connaissance sur les attributs discriminants de la classe. Sachant que l'importance des attributs varie selon les classes, chaque attribut aura une pondération spécifique à la classe traitée. L'exemple de l'attribut de forme illustre l'importance de ces connaissances. En effet, pour un objet de la classe "Végétation" la forme n'a pas d'importance car les différents types de végétation peuvent revêtir de nombreuses formes alors que pour la classe "Bâtiment", la forme reste un attribut discriminant vu que classiquement un "Bâtiment" est de forme rectangulaire. Ainsi, à chaque attribut est associé une pondération prise dans l'intervalle $[0, 1]$.

- Attributs qualitatifs. Cette catégorie comprend des informations qualitatives sur la classe. Nous y retrouvons par exemple les connaissances sur les relations des objets de cette classe avec d'autres objets. Par exemple, l'adjacence ou la proximité avec des objets d'autres classes [Derivaux, 2009, Forestier, 2010].

3.3 Template dynamique pour l'extraction des bâtiments

Dans cette section, nous introduisons une croissance sémantique hiérarchique pour l'extraction des bâtiments. Nous commençons par présenter une panoplie de méthodes objets pour l'extraction de bâtiments afin de situer notre approche d'extraction de bâtiments basée sur les templates dynamiques.

3.3.1 Méthodes objets pour l'extraction des bâtiments

Depuis plus de 20 ans, l'extraction des bâtiments à partir d'images numériques présente une des tâches de vision par ordinateur les plus complexes. Dans ce sens, plusieurs travaux ont été proposés dans la littérature [Jaynes *et al.*, 1994, Lari et Ebadi, 2007, Sledge *et al.*, 2011, Stankov et He, 2013]. Les approches orientées objets sont parmi les approches les plus prometteuses et les plus récemment utilisées [Attarzadeh et Momeni, 2012]. Ces approches sont basées sur l'architecture des approches d'analyse basées objets. La première étape est l'extraction des objets composant l'image. Ces objets sont décrits par un ensemble d'attributs utilisé pour identifier, dans la dernière étape, les objets bâtiments.

À partir d'une fusion entre des images panchromatiques et multi spectrales, Lari et Ebadi [Lari et Ebadi, 2007] proposent une approche de détection de bâtiment basée sur un ensemble de caractéristiques qu'ils jugent pertinent et discriminant par rapport aux autres classes de l'image. Dans un premier temps, un algorithme de croissance de région est appliqué pour extraire les zones homogènes. Ensuite, les objets extraits sont caractérisés par des attributs géométriques spécifiques (par exemple la surface et le périmètre), des attributs de structure (la compacité par exemple) et des attributs photométriques (moyenne d'intensité et couleur R, G, B). Finalement, des réseaux de neurone sont utilisés pour détecter le jeu de paramètres le plus discriminant pour la

Tableau 3.1 — Description des bâtiments (Dictionnaire de données Fodomust)

Type d'attribut	Attribut	Intervalle	
Spectral		min	max
	bleu	21.7	62.3
	vert	19.4	80.1
	rouge	29.7	135.1
	PIR	34.8	139
	NDVI	50.2	108
Spatial	Surface (m^2)	100	1000
	Largeur (m)	10	30
	Longueur (m)	10	30

caractérisation des bâtiments.

Aminipouri *et al.* [Aminipouri *et al.*, 2009] commencent par extraire les objets d'intérêts de l'image en utilisant une segmentation multi-résolution. Cette segmentation est réalisée à l'aide du logiciel *eCognition*. La classification est ensuite appliquée en utilisant une estimation des fonctions d'appartenance ensuite un classifieur par plus proches voisins.

Singh *et al.* [Singh *et al.*, 2012] utilisent une segmentation basée sur la mesure NDVI et des opérateurs morphologiques. Ensuite, afin de séparer les routes des bâtiments, les auteurs se basent sur des informations morphologiques.

Les limites de ces approches se joignent aux limites des approches d'analyse orientées objet vu qu'elles dépendent de la phase de construction des objets.

3.3.2 Template dynamique pour l'extraction de la zone de croissance

L'ensemble des connaissances qu'offre le dictionnaire des données présente une source importante d'information pouvant permettre une meilleure description des objets d'intérêts constituant l'image. Le problème majeur concernant ces connaissances reste leur intégration dans un processus d'identification d'objets géographique. Pour

les bâtiments, la description donnée concerne deux types d'attributs spécifiques, les attributs spectraux et les attributs spatiaux. Le tableau 3.1 présente une panoplie d'attributs décrivant la classe "Bâtiment".

Afin de profiter de ces connaissances dans le processus de croissance de régions sémantique, nous proposons de les intégrer dans la phase de croissance sémantique de la hiérarchie de croissance. Ces connaissances seront utilisées dans l'extraction de la zone de croissance pour chaque germe puis dans la phase de croissance de ce dernier.

L'une des propriétés fondamentales des bâtiments est qu'ils ont généralement une forme constante. Cette forme est généralement rectangulaire ou carré. Les figures 3.1 et 3.3 montrent deux représentations de la classe "Bâtiment" à savoir les pavillons et les bâtiments de l'habitat collectif. Dans le dictionnaire, les bâtiments d'habitat collectif sont décrits comme étant "des bâtiments de forme rectangulaire (étroit et allongé au sol) ou de forme carré de volume allongé vers le haut". Cette connaissance peut être primordiale afin de délimiter la zone de croissance d'un germe de la classe bâtiment.

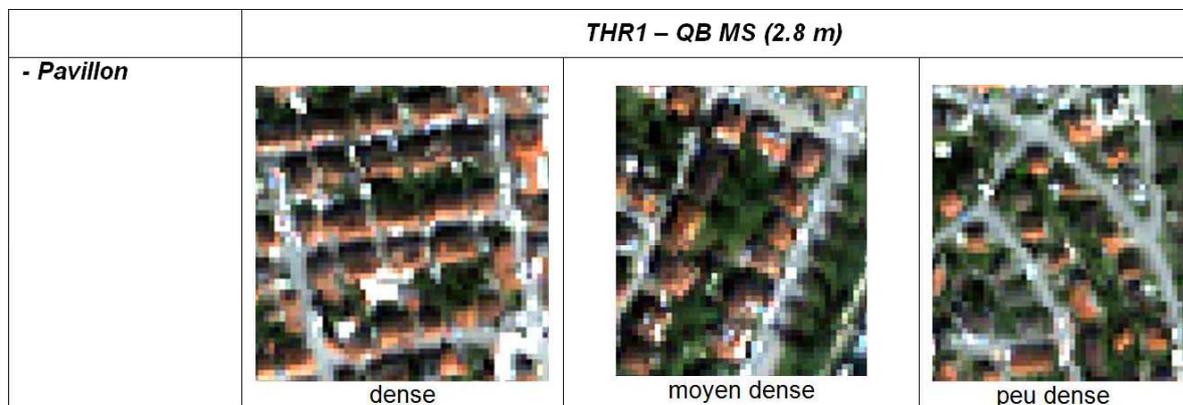


Figure 3.3 — Description visuelle des pavillons.

Une autre connaissance importante sur la classe "Bâtiment" concerne les propriétés géométriques de la forme. En effet, les bâtiments ont une superficie au sol variant entre une valeur minimale et maximale selon les règles de la ville où elle se situe. Pour la ville de Strasbourg par exemple, le tableau 3.1 illustre les dimensions que peut prendre un bâtiment. En nous basant sur ces données, nous remarquons que la classe "Bâtiment"

présente des propriétés spatiales intéressantes et qui peuvent permettre de limiter la zone de croissance. C'est pourquoi nous proposons d'utiliser un template construit à partir de ces connaissances de forme de longueur et de largeur pour extraire les régions pertinentes candidates lors de la phase de croissance pour chaque germe.

L'utilisation des templates pour l'extraction de bâtiments a été proposée dans plusieurs approches [Tinghua *et al.*, 2013]. Cependant, ces approches utilisent ces templates pour effectuer un appariement avec les objets détectés dans l'image afin d'en identifier les bâtiments. De même, pour les approches utilisant les processus ponctuels marqués [Ortner, 2004] dont le cadre stochastique est utilisé pour modéliser l'image par un ensemble de rectangles. Cependant, notre utilisation des templates n'a pas pour but d'extraire les bâtiments mais plutôt de délimiter la zone de croissance.

Afin de tester toutes les solutions possibles, nous simulons toutes les positions que peut prendre le bâtiment. Pour ce faire, nous projetons le template sur tous les angles entre 0 et 360° afin de détecter les régions candidates pour le processus de croissance. Le germe sera le centre de cette projection. La figure 3.4 illustre l'ensemble de l'espace de recherche généré par la projection sur les 360° . Les régions candidates seront donc l'ensemble des objets contenus dans les templates.

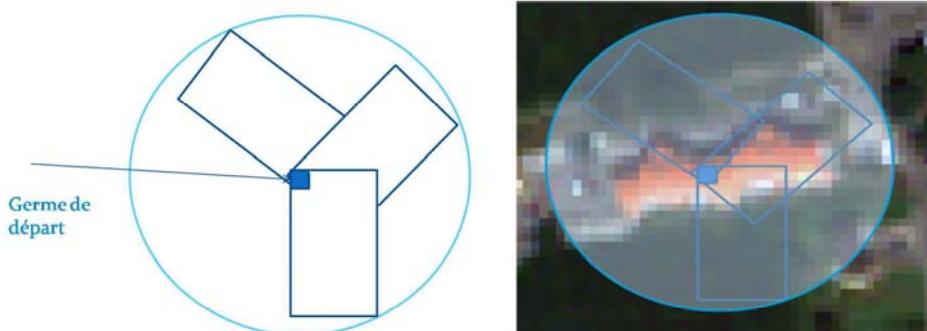


Figure 3.4 — Création du template.

Cette première projection permet donc de générer l'ensemble des candidats mais elle fait abstraction des connaissances acquises le long de toute la hiérarchie de croissance. En effet, les objets déjà traités lors des itérations précédentes ne peuvent pas être modifiés par des germes ayant une confiance inférieure. De même pour les germes de

même niveau de la hiérarchie. Ceci implique que ces régions doivent être intégrées dans le processus d'extraction de la zone de croissance comme des contraintes pour la sélection des régions candidates. La figure 3.5 illustre l'élagage des templates qui ne valident pas les contraintes. Les régions couvertes par le template en rouge ne seront pas intégrées dans la liste des candidats puisque celui-ci contient la région contrainte verte.

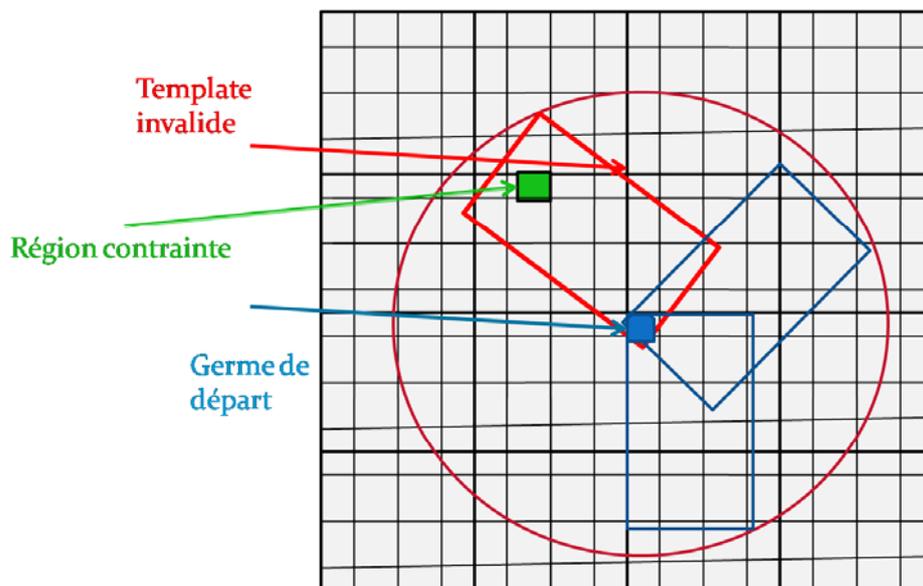
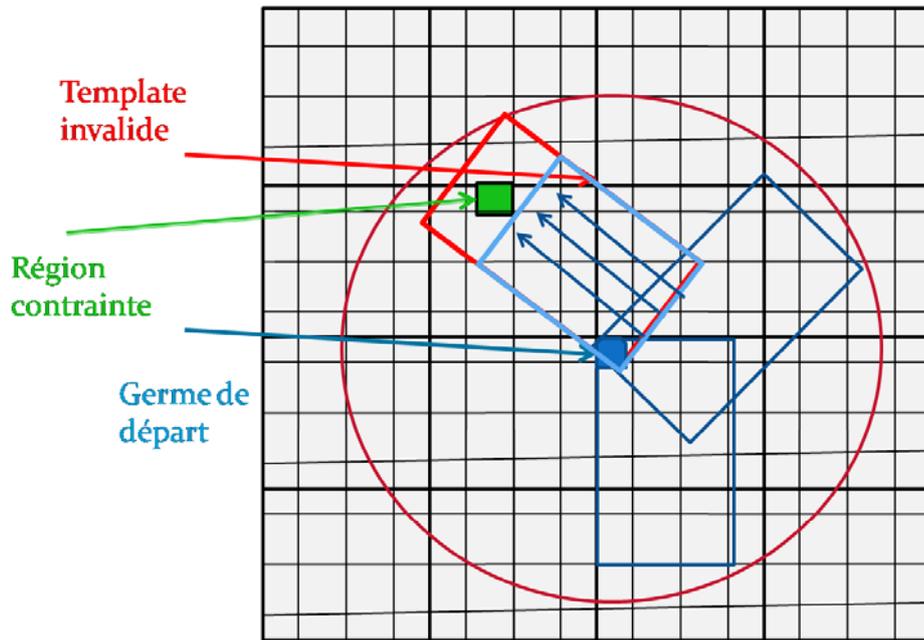
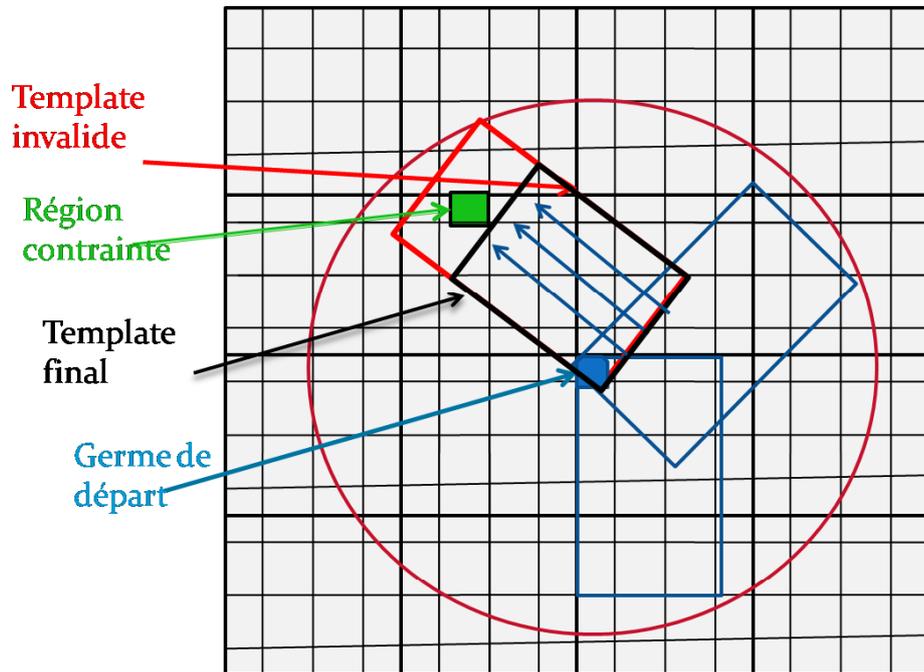


Figure 3.5 — Illustration des templates élagués.

Cet élagage complet des régions contenues dans les templates non validés peut provoquer la perte de régions candidates pertinentes pour la croissance. En effet, les régions entre le centre du template et les contraintes peuvent représenter des régions de la classe "Bâtiment". Afin d'y remédier, nous proposons de générer un template dynamique qui sera créé en prenant en compte les régions contraintes. Pour chaque angle, et au lieu de traiter tout le template, du centre du germe, nous faisons croître le template jusqu'à ce qu'il atteigne sa taille maximale ou qu'il rencontre une région contrainte. Les régions candidates seront donc les régions contenues dans le template généré. La figure 3.6 illustre le processus de création du template dynamique.



(a) Croissance du template



(b) Template final

Figure 3.6 — Création du template dynamique.

3.3.3 Croissance sémantique pour l'extraction des bâtiments

La croissance représente l'étape finale d'extraction de l'objet. En effet, après avoir extrait la zone de croissance du germe, ce dernier commence à fusionner avec l'ensemble de son voisinage en se basant sur un ensemble de critères. Cette croissance est limitée par la zone extraite dans l'étape précédente et un ensemble de connaissances. Afin d'intégrer les connaissances expertes dans le processus de croissance, nous proposons une fonction de fitness qui permet d'évaluer si la fusion entre le germe et son voisinage permet d'améliorer le résultat de l'extraction ou non. Cette fonction est fondée sur les connaissances expertes spécifiques à la classe "bâtiment".

Afin de sélectionner les régions à fusionner, nous utilisons la technique du meilleur ajustement [Lang *et al.*, 2009] en évaluant pour chaque voisin la région créée en fusionnant ce voisin avec le germe traité. Ensuite, les régions traitées sont triées par ordre décroissant. Si la valeur maximale de la fonction de fitness est supérieure à la fonction de fitness du germe, la région est fusionnée. La figure 3.7 illustre le fonctionnement d'une itération de la croissance d'un germe. En effet, pour le germe M , nous calculons la fonction de fitness spécifique à l'ensemble des régions de la fusion de ce germe avec ses régions voisines. La fusion de ce germe avec la région k est une région dont la fonction de fitness est maximale et atteint la valeur de 0.95. Cette valeur est aussi supérieure à la fonction de fitness du germe. Ce germe est donc fusionné avec la région K . Ce traitement est réitéré jusqu'à ce qu'aucune fusion du germe avec l'une de ces régions voisines n'accroisse la fonction de fitness à optimiser.

Pour définir cette fonction de fitness (cf. équation 3.1), nous proposons trois critères que nous pensons assez discriminants pour les bâtiments, à savoir, la forme, la taille et l'homogénéité.

$$\mathbb{F}(r) = \mathbf{Aire} \times \mathbf{R}_{GR} \times \mathbf{Hom}. \quad (3.1)$$

3.3.3.1 Aire

Les bâtiments ont une aire limite qu'ils ne peuvent pas dépasser. Afin d'évaluer cet intervalle, nous proposons une fonction qui varie entre 0 et 1. Cette fonction évalue la région traitée selon sa distance avec l'intervalle $I = [min_{aire}, max_{aire}]$ où min_{aire} (resp.

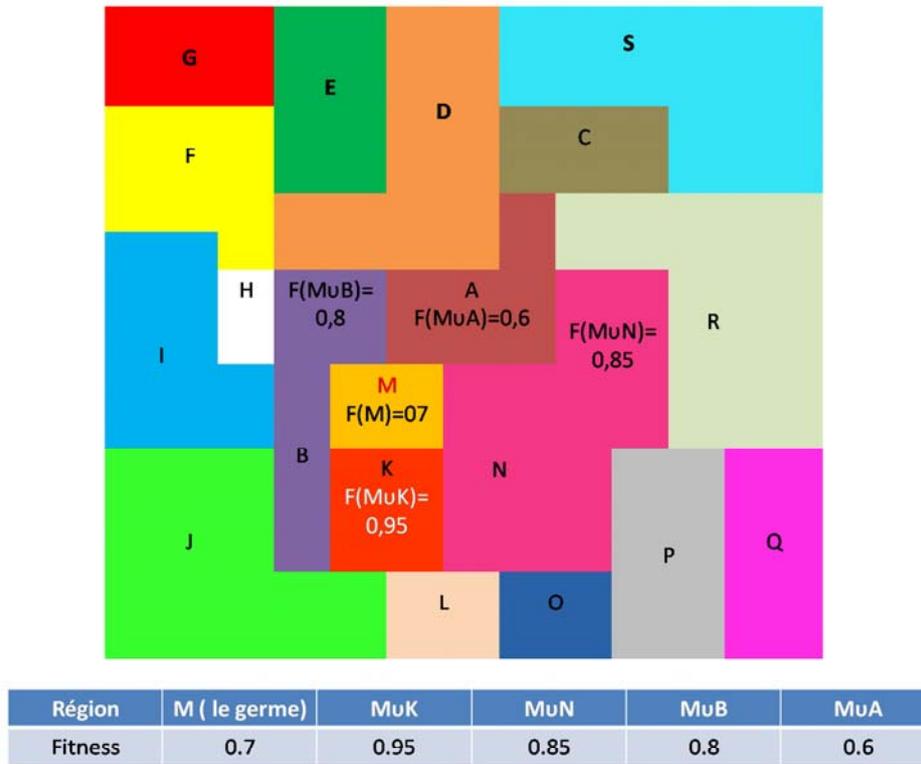


Figure 3.7 — Croissance par meilleur ajustement.

max_{aire}) représente la valeur minimale (resp. maximale) de l'aire d'un bâtiment. Cette fonction devra permettre de catalyser ou non la croissance de la région tant qu'elle ne dépasse pas la borne supérieure max_{aire} .

Soit r la région à évaluer. La fonction d'évaluation de l'aire $Aire$ est définie de la manière suivante :

$$Aire = \begin{cases} 1 & \text{si } aire(r) \in [min_{aire}, max_{aire}] \\ \frac{aire(r)}{min_{aire}} & \text{if } aire(r) < min_{aire} \\ \frac{max_{aire}}{aire(r)} & \text{if } aire(r) > max_{aire} \end{cases} \quad (3.2)$$

Cette fonction permet de catalyser la croissance si la taille de l'objet qui est entrain de croître est inférieure à la borne inférieure. En effet, dans l'intervalle $[0, min_{aire}[$ la fonction est croissante. Si l'aire de la région est dans l'intervalle $[min_{aire}, max_{aire}]$, la fonction est alors constante et elle n'a plus d'effet sur l' $Aire$ vu que sa valeur se stabilise à 1 permettant ainsi aux autres attributs d'être plus discriminants. Lorsque cette valeur dépasse la borne supérieure (i.e. dans l'intervalle $]max_{aire}, +\infty[$), cette

fonction devient décroissante et elle pénalise la croissance. La figure 3.8 illustre l'allure de la courbe.

3.3.3.2 Rectangularité

La forme rectangulaire peut être une propriété utile et appliquée à une gamme de tâches diverses telles que la détection des routes, la détection des bâtiments ou le contrôle de qualité en industrie. Dans la littérature, nous trouvons plusieurs méthodes de mesure de rectangularité. La méthode du rectangle minimal englobant est considérée comme la méthode standard. Cette métrique, qui calcule le cadre de sélection de la forme de l'objet, est sensible au bruit.

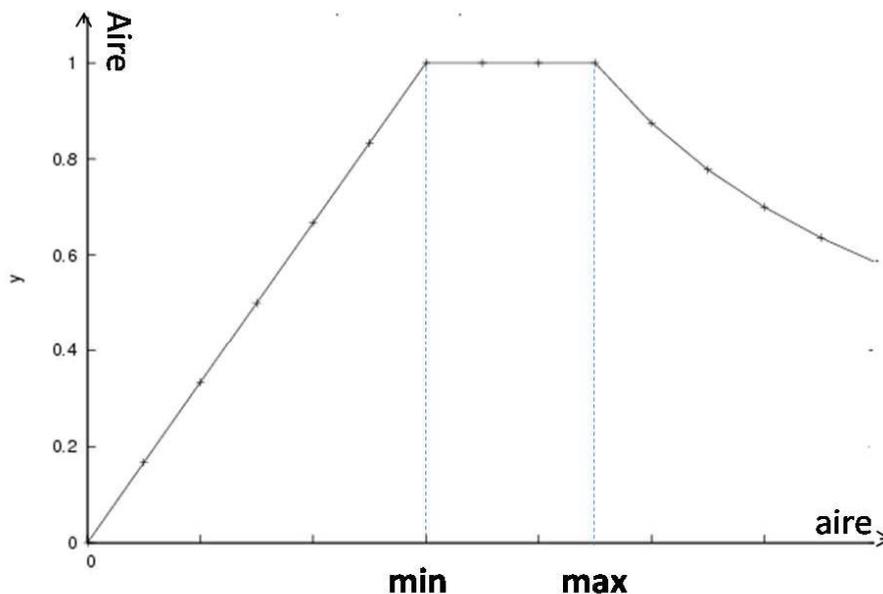


Figure 3.8 — Allure de la courbe de la fonction *Aire*.

Rosin a essayé de surmonter ce problème en introduisant trois nouvelles méthodes [ROSIN, 1999] :

- La méthode de l'accord (RA) assume que la forme est rectangulaire et estime les dimensions de la région par deux méthodes différentes. La forme rectangulaire

est définie comme le degré d'accord entre ces deux estimations. Cette métrique reste très sensible au bruit.

- La méthode des moments (RM) caractérise la forme en se basant sur son moment. Cette métrique peut conduire à une estimation incorrecte de la rectangULARITÉ puisqu'elle est sensible aux transformations géométriques de base (rotation, translation et changement d'échelle).
- La méthode de différence (RD) calcule l'enveloppe de la région mais en utilisant son image ellipse qui ajuste mieux la région. Cette métrique est moins sensible aux saillies mais non invariante vis-à-vis des rotations et des translations.

D'autres recherches dans ce domaine sont développées afin de surmonter le problème de la sensibilité aux transformations géométriques appliquées sur un objet en se basant sur la transformée de Radon. Le principe consiste à comparer la signature d'un objet avec la signature théorique d'un modèle rectangulaire parfait.

Tabbone *et al.* [Tabbone *et al.*, 2008] proposent une métrique de rectangULARITÉ basée sur la *R* – signature et qui opère seulement sur les objets pleins. Par contre, Nacereddine *et al.* [Nacereddine *et al.*, 2010] proposent une méthode basée sur la Φ – signature et utilisent seulement les contours des objets traités. Pour les deux approches, l'objet de signature est comparé à une signature modèle d'un rectangle idéal. La décision de la rectangULARITÉ des objets selon ces deux signatures se base sur la comparaison des résultats de leur application sur l'objet étudié avec ceux du modèle statistique calculé au préalable (la signature d'un rectangle parfait). Il s'agit de mesurer leur taux de similarité sans spécifier une métrique de rectangULARITÉ.

De leur part, Hentati *et al.* [Hentati *et al.*, 2011] proposent une métrique de rectangULARITÉ, la R_{GR} , simple et efficace qui utilise un nouveau descripteur de forme appelé la *GR* – signature (basée sur la transformée de Radon et le gradient) pour calculer le taux de rectangULARITÉ d'un objet donné. Une étude comparative [Hentati *et al.*, 2011] montre que la *GR* – signature permet une meilleure discrimination de la forme rectangulaire par rapport à la *R* – signature = et la Φ – signature. Pour cette raison, nous choisissons d'utiliser la métrique R_{GR} basée sur la *GR* – signature dans la phase de mesure de rectangULARITÉ des bâtiments extraits. Ce paragraphe explique brièvement leurs principes de fonctionnement.

Par définition, la transformée de Radon [Radon, 1917] d'une image $f(x, y)$ est déterminée par une série de projections de l'image le long des lignes radiales prises

sous des angles différents. Une présentation plus détaillée de la transformée de Radon sera introduite dans la section 3.4.2.2. L'application de la transformée de Radon sur f génère des pics qui présentent les variations brusques de la fonction d'intensité de la matrice R de Radon. Vu la capacité du gradient à caractériser les discontinuités d'une fonction, les auteurs l'appliquent sur la matrice R résultat de l'application de la transformée de Radon sur l'image initiale. D'où la création de la matrice GR (un signal 3D). Afin d'extraire une information réduite et pertinente (un signal 2D), ils ne conservent, pour chaque angle Θ , que les deux points ayant des valeurs extrêmes (le maximum et le minimum). D'où la définition de la $GR - signature$ d'un objet qui est constituée de deux signaux : $GRpositive$ et $GRnegative$ qui sont respectivement les maxima et les minima des colonnes de GR . Cette signature présente plusieurs propriétés intéressantes :

- périodicité de période 2π ;
- les deux signaux de la $GR - signature$, $GRpositive$ et $GRnegative$, sont de signes opposés et en opposition de phase (la différence de phase est π) ;
- invariance à la translation.

Étant donné que les deux signaux de la $GR - signature$ sont de signes opposés et en opposition de phase, la description complète de l'objet est obtenue soit par un seul 2π -signal ou par deux π -signaux. [Hentati *et al.*, 2011] adoptent une signature sur un intervalle π .

La métrique R_{GR} de mesure de rectangularité se base sur les caractéristiques du la $GR - signature$ des objets rectangulaires. La $GR - signature$ d'un rectangle parfait contient deux paires de pics opposés dont leur ordonnée "amplitude A" dépend des dimensions de l'objet et leur abscisse "angle Θ " dépend de son orientation. La différence d'angle entre le grand et le petit pic de chaque signal de la $GR - signature$ est 90° . Elle représente l'angle entre les deux médianes du rectangle, i.e., $\theta_{measure}$ et $A_{measure}$. Dans le cas non parfait, des différences en angle et en amplitude apparaissent entre les paires des pics (figure 3.9) et qui agissent sur le taux de rectangularité de l'objet étudié.

La formule de la R_{GR} est la suivante :

$$R_{GR} = \frac{\theta_{measure} + A_{measure}}{2} \quad (3.3)$$

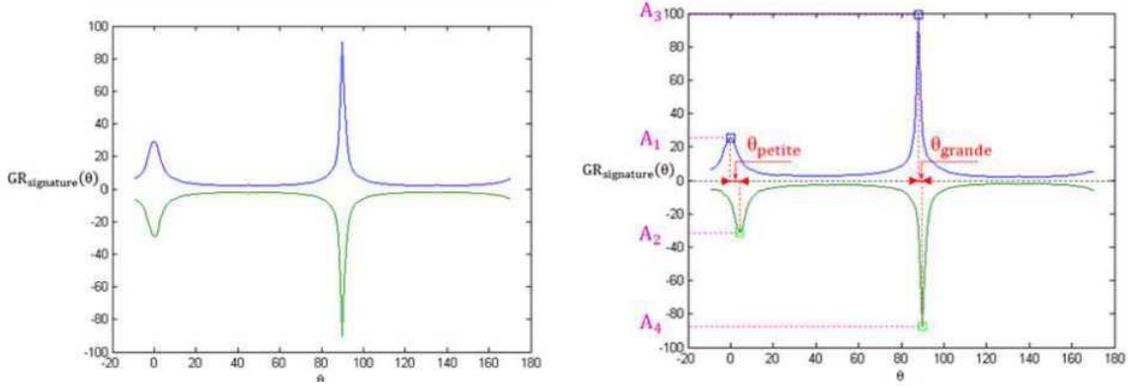

 (a) $GR - signature$: objet rectangulaire parfait (b) $GR - signature$: objet rectangulaire non parfait

Figure 3.9 — Différence entre la $GR - signature$ d'un objet rectangulaire parfait et la $GR - signature$ d'un objet rectangulaire non parfait.

Où :

$$\theta_{measure} = \frac{90 - (\theta_{petite} + \theta_{grande})}{90} \quad (3.4)$$

et

$$A_{measure} = 1 - \frac{A_2 - A_1}{A_4 - A_3} \quad (3.5)$$

Avec θ_{petite} , la différence en angle entre les deux petits pics et θ_{grande} , la différence en angle entre les deux grands pics. Les A_i sont les amplitudes normalisées des quatres pics selon la formule :

$$A_i = \frac{A_i - A_{min}}{A_{max} - A_{min}} \quad (3.6)$$

Ainsi, la R_{GR} est égale à 1 pour les rectangles parfaits et se rapproche de 0 pour les objets non rectangulaires.

3.3.3.3 Homogénéité

L'une des propriétés des bâtiments est leur homogénéité en terme de radiométrie. Par ailleurs, nous constatons que dans la plupart des cas, les bâtiments contrastent avec leur voisinage. Cette connaissance n'est pas toujours valable. Toutefois, compte tenu de la spécificité des bâtiments de la ville de Strasbourg qui impose certaines conditions de construction pour maintenir l'homogénéité du paysage urbain, cette hypothèse est admise dans la fonction de fitness.

L'homogénéité devra représenter alors le facteur qui indique le degré de similarité de point de vue radiométrie entre la région qui croit et l'ensemble des candidats à la croissance. Pour ce faire, nous introduisons l'équation 3.7.

$$\mathbf{Hom} = 1 - (|Diff_r| + |Diff_g| + |Diff_b| \times 0.01) \quad (3.7)$$

Où $Diff_i$ représente pour chaque band spectrale la différence entre la moyenne de la radiométrie de la région traitée et des régions candidates. Cette fonction est une fonction décroissante qui va favoriser la fusion des régions homogènes, plus la différence entre la région traitée et la région candidate augmente et plus la fonction tends vers 0. A un certain seuil, cette fonction devient négative ce qui implique l'élimination de la région candidate.

3.4 Extraction des routes basée sur la transformé de Radon

Dans cette section, nous introduisons notre proposition pour l'extraction des routes. Nous commençons tout d'abord par présenter quelques méthodes objets pour l'extraction des routes. Ensuite, nous introduisons notre solution.

3.4.1 Méthodes objets pour l'extraction des routes

L'extraction de réseaux routiers, à partir d'images satellitaires ou aériennes, a fait l'objet de nombreuses recherches et une grande variété de méthodes ont été proposées pour traiter ce problème. Les approches semi-automatiques telles que les contours actifs [Kass *et al.*, 1988, Rochery *et al.*, 2005] ou la programmation dynamique [Fischler *et al.*, 1981, Merlet et Zerubia, 1996] restent coûteuses et dépendantes de l'intervention humaine. Certes, le concept d'extraction des réseaux routiers est relativement simple, mais l'introduction de processus d'extraction fiables reste un défi difficile à relever. Il n'existe pas d'algorithme générique suffisamment fiable pour tous les types d'images. L'extraction de route reste en grande partie, un processus manuel coûteux. L'inconvénient majeur des approches semi-automatiques et des approches bas niveaux reste leur sensibilité au bruit, en particulier pour des images haute résolution dans laquelle les artefact inhérents à la scène observée sont ajoutée (par exemple l'ombre

des arbres sur les routes). Afin de réduire cette sensibilité au bruit, les approches basées objets ont été introduites.

Dans la dernière décennie, plusieurs approches d'extraction de réseaux routiers ont été proposées et spécialement pour l'analyse des images multi-spectrales [Zhang et Couloigner, 2006b, Zhou et Troy, 2008, Naouai *et al.*, 2010, Zhang *et al.*, 2010]. Les approches d'extraction de réseaux routiers sont composées de deux parties : (i) une bi-classification de l'image en deux classes route et non route. Cette classification est généralement basée sur les propriétés spectrales. (ii) La deuxième étape est la localisation des objets routes.

Zhang et Couloigner [Zhang et Couloigner, 2006b] introduisent une approche d'extraction des réseaux routiers basée objets. Ils commencent par une segmentation utilisant un algorithme de k-moyennes. Ensuite, les objets de la classe "route" sont identifiés en utilisant un classifieur flou basé sur un ensemble de fonctions d'appartenance. Les auteurs utilisent les propriétés texturales afin d'éviter les ambiguïtés avec les classes proches des routes telles que les parkings ou les toits des bâtiments. Ensuite, la transformée de Radon est utilisée afin d'extraire les squelettes des objets "route". Finalement, ces objets sont groupés afin d'obtenir le réseau final.

Zhao et Wang [Zhao et Wang, 2010] proposent un algorithme d'extraction de routes à partir des images haute résolution. Dans un premier temps, un algorithme de segmentation par seuillage simple tenant en compte de distribution de niveaux de gris des objets est appliqué sur l'image originale. L'élimination de bruit présent dans l'image binaire est réalisée par l'application d'une combinaison d'opérations de morphologie mathématique. Le processus morphologique est fondé sur plusieurs étapes. Durant la première étape le bruit présent dans la région de routes est filtré. Ensuite, la région résultante est déconnectée du bruit relatif au reste de l'image via une opération de fermeture de disque. Finalement, toutes les régions non routes sont supprimées. L'extraction de routes est effectuée en utilisant le modèle des contours actifs. L'algorithme est appliqué sur des images de télédétection haute résolution.

Naouai *et al.* [Naouai *et al.*, 2011] proposent une méthode basée objets fondée sur les règles de perception visuelle. Une première étape de vectorization permet le passage du pixel vers l'objet. Ensuite, un ensemble de règles inspirées de la théorie gestaltiste permet de fusionner les régions composants la route.

L'ensemble de ces approches reste dépendant de la phase de localisation des régions

et donc de la phase de segmentation initiale. Dans ce cadre, nous introduisons notre approche de croissance de région hiérarchique basée sur la transformée de Radon.

3.4.2 Transformée de Radon

3.4.2.1 Transformée de Radon continue

La transformée de Radon classique porte le nom de son créateur Johann Radon. Il a permis avec la transformée de Radon classique la détermination d'une fonction à travers des valeurs de ses intégrales le long des lignes du plan [Radon, 1917]. Plusieurs auteurs se sont mis d'accord sur le fait que la transformée de Hough [Hough, 1962] est un cas particulier de la transformée de Radon. La définition formelle de la transformée de Radon classique est la suivante :

Soit $f(x, y)$ une image définie dans l'espace euclidien de deux dimensions et $T_R f(\rho, \theta)$ sa transformée dans le domaine Radon :

$$T_R f(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy. \quad (3.8)$$

avec $\theta \in [0, \pi]$, $\rho \in]-\infty, \infty[$ et $\delta(\cdot)$ la fonction de Dirac. L'espace (ρ, θ) est généralement appelé l'espace Radon (ER) ou l'espace des paramètres. L'équation $\rho = x \cos(\theta) + y \sin(\theta)$ représente l'équation paramétrique d'une ligne droite en coordonnées polaires. L'équation. 3.8 représente l'intégration de l'image le long de chacune de ces lignes droites et le stockage de la valeur résultante dans le point de l'ER correspondant. La figure 3.10 illustre l'acquisition d'une projection. Sachant que l'image n'est pas représentée dans le domaine continu, une adaptation pour le domaine discret a été proposée.

3.4.2.2 Transformée de Radon discrète

Soit un domaine discret $2D$ noté I dans lequel la fonction continue f est échantillonnée. f est plongée dans une grille régulière discrète (à gauche sur la figure 3.11). L'équation 3.9 définit l'étape d'échantillonnage [Toft, 1996]. Chaque fonction $f(x, y)$ est associé à la valeur du pixel (i, j) qui le contient. Autrement dit, tous les points (x, y) d'un même pixel ont la même valeur. L'image à droite sur la figure 3.11

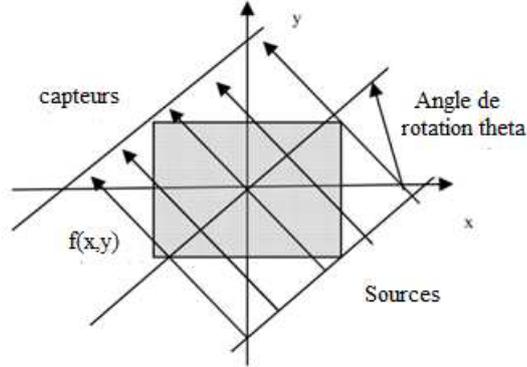


Figure 3.10 — Projection selon des rayons parallèles à l'angle θ_0 [Wang et Li, 2006].

illustre ce processus.

$$f(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I(i, j) \Delta(x - j) \Delta(y - i) \quad (3.9)$$

avec $\Delta(\cdot)$ l'impulsion de Dirac discrète.

Les équations 3.8 et 3.9 [Toft, 1996] permettent de définir la transformée de Radon discrète comme la somme des valeurs des pixels traversés par la droite dans l'image.

Cette valeur est notée $T_{RD}I(\rho, \theta)$ et est donnée par la formule suivante :

$$T_{RD}I(\rho, \theta) = \sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} I(i, j) \times \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Delta(x - j) \Delta(y - j) \delta(\rho - x \cos \theta - y \sin \theta) dx dy \right) \quad (3.10)$$

La transformée de Radon discrète (équation 3.11) calcule la valeur de projection

$T_{RD}I(\rho, \theta)$ depuis une image I de taille $L \times H$ pixels :

$$T_{RD}I(\rho, \theta) = \sum_{i=0}^{L-1} \sum_{j=0}^{H-1} I(i, j) \Delta(\rho - j \cos \theta - i \sin \theta) \quad (3.11)$$

avec $\Delta(\cdot)$ l'impulsion de Dirac discrète qui vaut 1 si la droite (θ, ρ) traverse le pixel (i, j) et 0 sinon.

3.4.3 La transformée de Radon pour l'estimation de la direction des routes

La classe "route" ne possède pas les mêmes propriétés de la classe "Bâtiment". En effet, les routes présentent dans les images optiques une radiométrie plus homogène

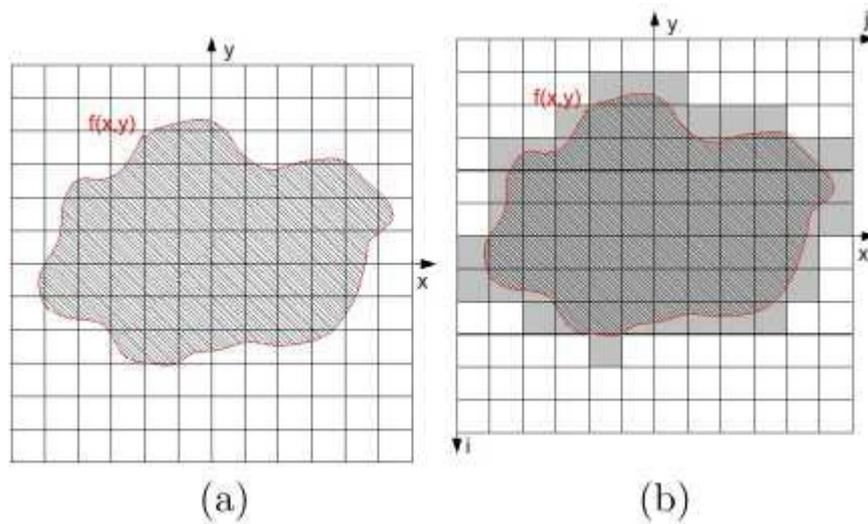


Figure 3.11 — Plongement du domaine f dans une grille discrète.

que les bâtiments. Elles sont généralement recouvertes d'asphalte. Les routes ont des contours linéaires et parallèles, sont généralement au niveau du sol et ne projettent pas d'ombre. La plupart des routes ne contiennent pas de végétation [Poulain, 2010]. Le dictionnaire des données décrit les routes par un ensemble de propriétés spatiales et spectrales. En effet, il présente les routes comme des objets homogènes ayant une largeur entre 6 et 25 m. Afin d'intégrer la notion de largeur, il est nécessaire de détecter la direction de la route pour chaque germe traité afin de pouvoir limiter l'espace de recherche. Une des représentations graphiques des routes est illustrée dans la figure 3.12.

	THR1	HR1 (15 m)	HR2 (30 m)
- autre route		NON	NON

Figure 3.12 — Illustration graphique de la classe "Route", extrait du dictionnaire des données.

Nous proposons d'utiliser la transformée de Radon pour détecter la direction des routes. Même si l'utilisation de la transformée de Radon est commune pour l'extraction des réseaux routiers [Zhang et Couloigner, 2006a, Silva *et al.*, 2010], les travaux cités restent focalisés sur la détection des objets de forme linéique dans l'image et qui seront assimilés à des portions de routes. La solution que nous proposons s'oriente plus vers la détection de la direction de la route à partir d'une région germe de la classe route. Pour ce faire, nous proposons de profiter de la nature hiérarchique de notre approche afin d'utiliser les objets déjà détectés et des germes de même nature (i.e. des germes de la classe route).

L'idée est de détecter la direction qui intersecte le plus grand nombre d'objets de la classe "route". Pour chaque germe, nous commençons par extraire une fenêtre centrée sur l'objet traité. Cette fenêtre ne contient que les objets de la classe "route" qui sont déjà détectés ou les germes de cette classe. La figure 3.13 présente un exemple d'extraction de la fenêtre centrée sur le germe en bleu.

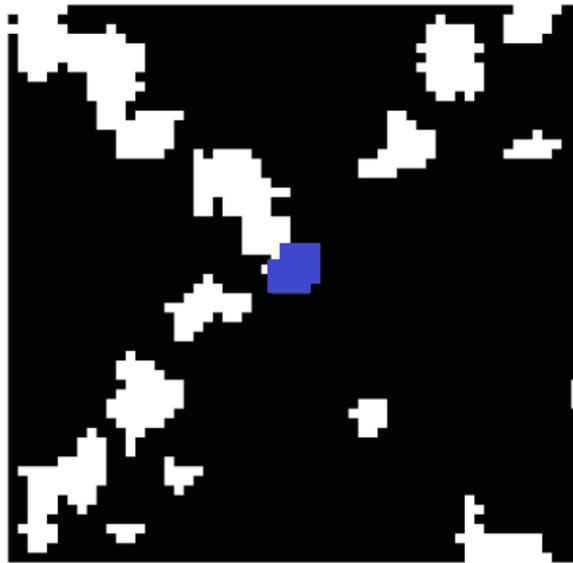


Figure 3.13 — Extraction de la fenêtre centrée sur le germe pour l'application de la transformée de Radon.

À partir de cette fenêtre, nous appliquons la transformée de Radon. Le résultat sera l'ensemble des 180 projections qui représentent les 180 angles de projections utilisées par la transformée de Radon. Le résultat de l'application de la transformée de Radon est illustré par la figure 3.14.

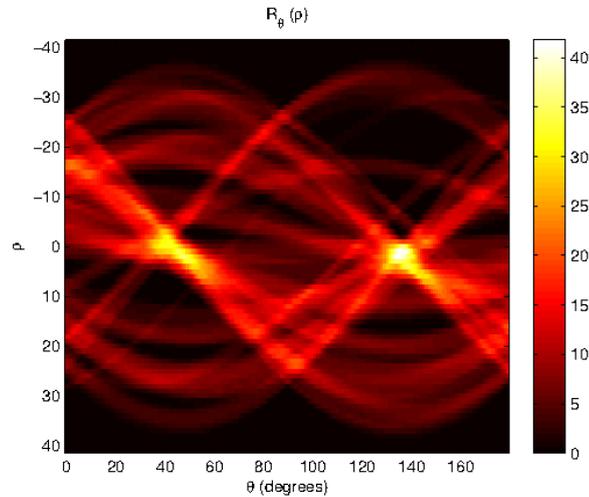
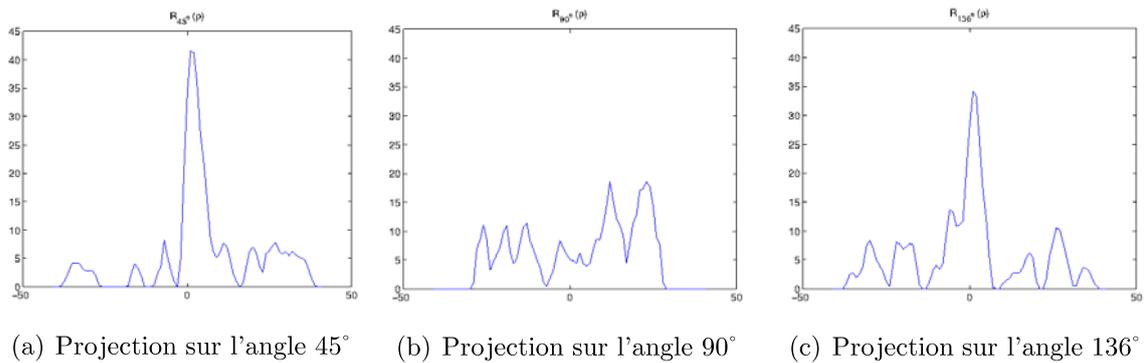


Figure 3.14 — Résultat de l'application de la transformée de Radon.



(a) Projection sur l'angle 45° (b) Projection sur l'angle 90° (c) Projection sur l'angle 136°

Figure 3.15 — Résultat de la projection sur trois angles différents.

Les deux points lumineux représentent les deux pics qui sont les directions des projections qui contiennent le plus d'objets "route". La figure 3.15 présente 3 projections selon 3 angles différents. La direction extraite est celle qui présente le pic le plus long. C'est la projection sur l'angle 45° (figure 3.15(a)).

Cette direction est illustrée par la ligne bleu sur la figure 3.16. Nous remarquons que c'est réellement la direction qui contient le plus d'objets de la classe " route ".

3.4.4 Extraction des routes

La phase d'extraction des routes consiste en une croissance à partir du germe de départ en nous basant sur la direction extraite en utilisant la transformée de Radon. Une deuxième connaissance peut être utilisée afin d'extraire la zone de croissance. En

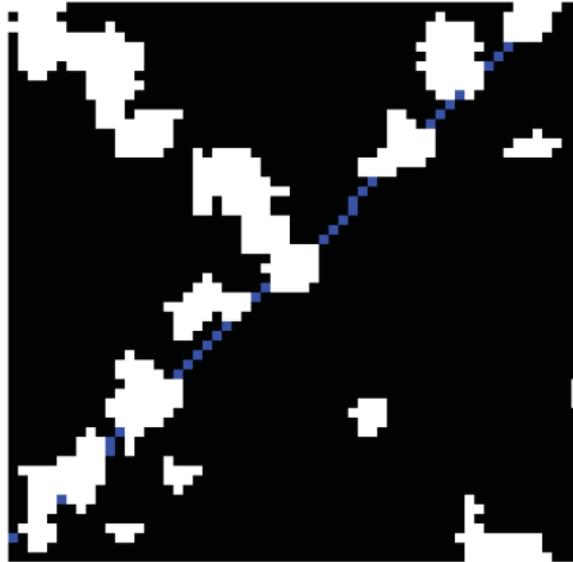


Figure 3.16 — La direction détectée par l'application de la transformée de Radon.

effet, les routes ont une largeur qui ne peut pas dépasser un seuil maximal que l'on notera L_{max} . En se basant sur la direction du germe ainsi que sur la largeur L_{max} , nous pouvons restreindre l'espace de croissance du germe. Ne connaissant pas la position du germe par rapport à la route, l'estimation de la largeur devra se baser sur le pire des cas, à savoir, que le germe puisse se placer sur les deux extrémités de la route. La figure 3.17 illustre l'extraction de la zone de croissance sur le germe présenté dans la figure 3.13. Les deux zones de croissance inférieure et supérieure permettent de créer la zone de croissance finale. Ces deux limites de couleur verte et orange ont les deux la même largeur L_{max} et sont parallèles à la direction de la route extraite en appliquant la transformée de Radon. La limite finale de la zone de croissance est la zone en tiret rouge qui est l'union des deux zones inférieure et supérieure.

La solution proposée ne se restreint pas à la forme linéaire. En effet, l'utilisation de fenêtre centrée sur les germes traités permet d'assimiler les routes de forme circulaire et les routes qui présentent des oscillations à des routes linéaires par morceaux. La figure 3.18 illustre le morcellement des routes de forme circulaire et les routes présentant des oscillations. Les lignes en traits discontinus présentent la vision linéaire par morceaux des deux formes en bleu.

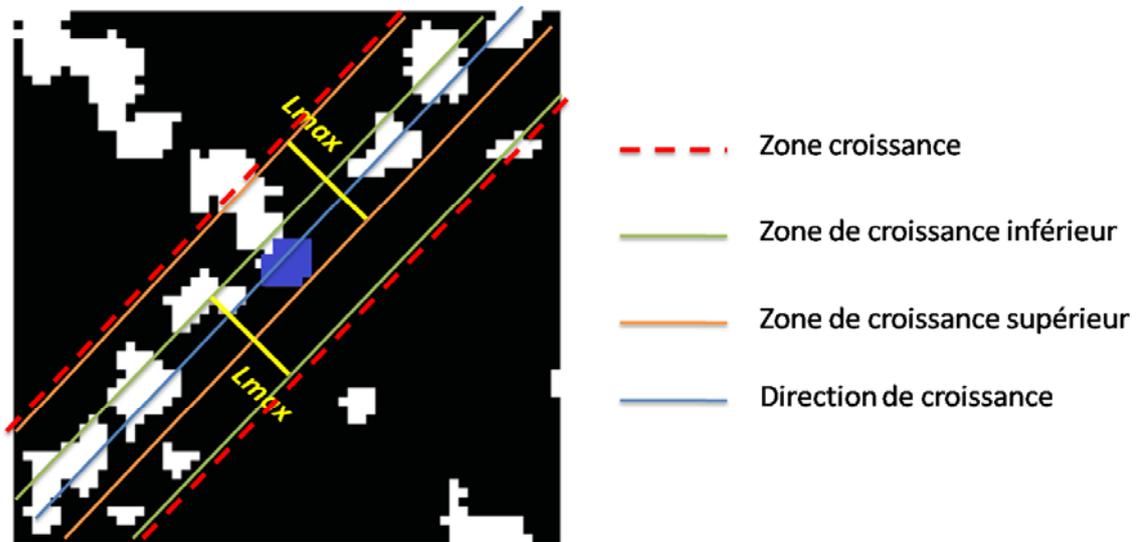


Figure 3.17 — Extraction de la zone de croissance pour un germe de type route.

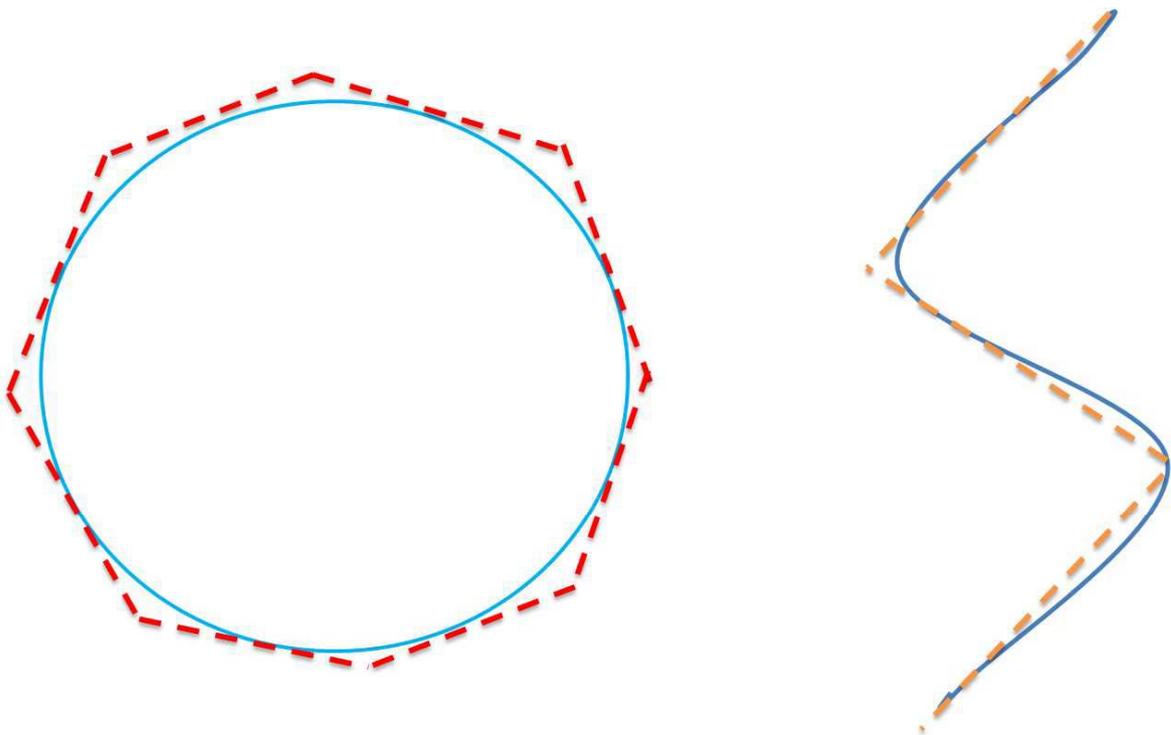


Figure 3.18 — Routes linéaires par morceaux.

Une fois la zone de croissance extraite, nous entamons la phase de croissance à partir du germe. Les routes ne présentant pas de propriétés de forme, à partir de cette zone, la croissance se fait en se basant sur le critère d'homogénéité. La technique de

sélection est la même que celle utilisée pour les bâtiments qui est la technique du meilleur ajustement.

3.5 Expérimentation et évaluation

Cette section introduit les résultats expérimentaux relatifs à l'approche proposée. Nous présentons d'abord les paramètres d'expérimentation utilisés pour ensuite discuter les résultats obtenus. Dans le chapitre 6, nous présentons une expérimentation plus poussée de notre approche.

Afin d'évaluer la performance d'un système, il est nécessaire de définir des critères permettant d'estimer la pertinence de ce dernier. Les mesures les plus utilisées pour l'évaluation sont l'exactitude et la pertinence des résultats, le temps de réponse du système ainsi que l'utilisation de la mémoire. Cependant, dans le contexte visé par nos recherches, à savoir, l'extraction d'objets urbains dans des images à très hautes résolution, nous nous focalisons plus sur l'aspect qualitatif des résultats. Pour ce faire, nous présentons ici les deux mesures les plus utilisées dans la littérature et qui sont le rappel et la précision.

Le rappel représente le nombre d'individus détectés par rapport à l'ensemble des individus détectés. Dans le contexte de nos travaux, un individu représente un objet géographique i.e., route, bâtiment. Cette mesure permet donc de mesurer la capacité du système à restituer l'ensemble des objets pertinents. Le rappel est calculé par la formule suivante :

$$Rappel = \frac{\text{Nombre d'objets correctement identifiés}}{\text{Nombre d'objets pertinents dans l'image}} \quad (3.12)$$

La précision représente le nombre d'objets correctement identifiés par rapport au nombre d'objets identifiés. Il permet d'évaluer la capacité du système à ne restituer que des objets pertinents. La précision est calculée par la formule suivante :

$$Precision = \frac{\text{Nombre d'objets correctement identifiés}}{\text{Nombre d'objets identifiés}} \quad (3.13)$$

Nous présentons une évaluation de notre approche sur les deux images suivantes 3.19(a) et 3.19(b) représentant respectivement une zone éparse et une zone dense de la

Tableau 3.2 — Évaluation du rappel et de la précision

Mesures/Classes	Route	Bâtiment
Rappel	0.8735	0.9424
Précision	0.8528	0.8740

zone urbaine de Strasbourg.



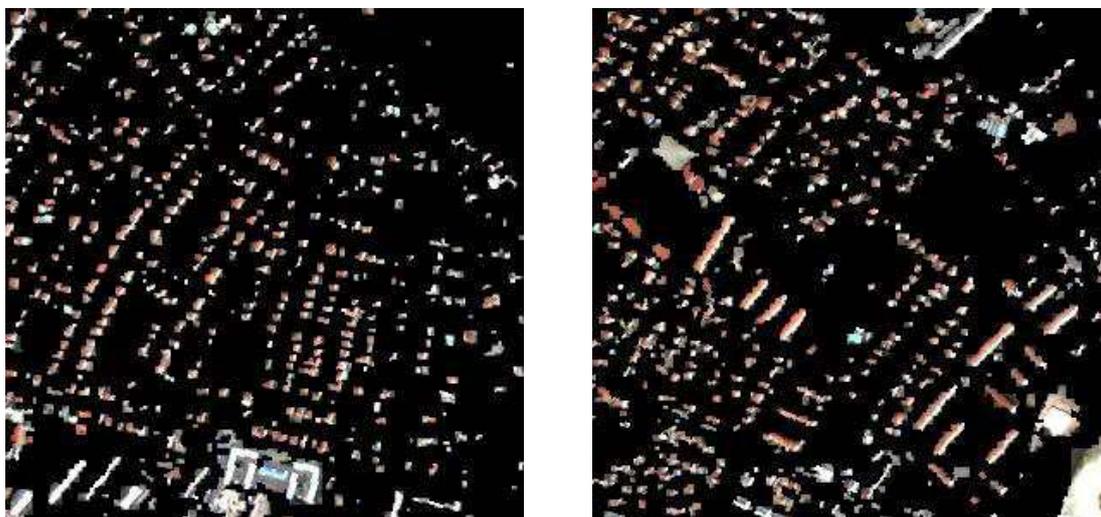
(a) Zone éparse

(b) Zone dense

Figure 3.19 — Zones urbaines de la ville de Strasbourg

Nous présentons ici le résultat de l'extraction concernant les classes "Route" et "Bâtiment". Les figures 3.20(a) et 3.20(b) illustrent le résultat d'extraction des bâtiments pour la zone de test éparse et dense alors que les résultats de la classe route sont illustrés dans les figures 3.21(a) et 3.21(b). Nous remarquons que la majorité des bâtiments et des routes présents dans l'image ont été extraits et que la qualité de l'extraction s'avère acceptable pour les images denses et éparsees. Cependant, les résultats des zones éparsees sont meilleurs. Ceci est prévisible puisque les zones denses présentent plus de bruit.

L'évaluation numérique confirme l'intérêt de notre approche. Le tableau 3.2 détaille le résultat des mesures de rappel et de précision pour les deux classes route et bâtiment. Le rappel varie entre 87,35% et 94,24% alors que la précision varie entre 85,28% et 87,40%. Les valeurs de la précision sont inférieures à celles du rappel à cause des erreurs de classification de la phase de prétraitement.



(a) Zone éparse

(b) zone dense

Figure 3.20 — Extraction de la classe "Bâtiment"



(a) Zone éparse

(b) Zone dense

Figure 3.21 — Extraction de la classe "Route"

3.6 Conclusion

Au cours de ce chapitre, nous avons proposé une application de notre approche de croissance sémantique hiérarchique sur les deux classes les plus traitées pour l'analyse des zones urbaines qui sont la classe "Route" et la classe "Bâtiment". Pour les deux classes, nous avons intégré un ensemble de connaissances expertes extraites à partir du dictionnaire des données créé lors du projet Fodomust.

Pour la classe "Bâtiment", nous avons utilisé les propriétés géométriques des bâtiments afin de délimiter la zone de croissance. Pour la croissance, nous avons aussi utilisé les connaissances géométriques auxquelles nous avons ajouté les connaissances radiométriques.

Pour la classe "Route", nous avons utilisé la transformée de Radon pour extraire la direction de chaque germe. À partir de cette direction, nous estimons la zone de croissance et nous y appliquons une croissance basée sur l'homogénéité radiométrique.

Après avoir présenté, dans les deux précédents chapitres, notre première approche basée sur les techniques de segmentation région, nous introduisons dans le chapitre suivant notre deuxième approche basée sur les techniques de collaboration région-contour. L'approche proposée permet une collaboration entre les techniques de segmentation région et contour dans le cadre d'un système multi-agent. Une vision locale est introduite en traitant chaque objet d'intérêt séparément permettant une meilleure adaptation du paramétrage des agents utilisés.

4

Systeme multi-agents évolutionnaire pour une approche coopérative régions-contours pour l'extraction des objets d'intérêt

4.1 Introduction

Comme nous l'avons introduit dans le premier chapitre, les approches que nous proposons se basent sur les techniques de segmentation les plus propices à la coopération. La première approche que nous avons proposée est basée sur les techniques de croissance de région. Dans ce chapitre, nous allons introduire notre deuxième approche collaborative basée sur une collaboration région contour. En effet, les deux grandes familles de segmentation les plus utilisées sont les approches contours et régions. Même si elles présentent de bons résultats d'extraction, ces approches présentent aussi certains inconvénients. Les approches contours peuvent présenter de faux positifs dûs à la sensibilité au bruit des opérateurs utilisés. Un deuxième inconvénient majeur est le problème des contours qui sont généralement non connectés. Les approches régions quant à elles, présentent certes des contours fermés, mais qui ne coïncident généralement pas avec les objets d'intérêt. Cet inconvénient résulte du problème de paramétrage et de condition de croissance que présentent ces approches. Enfin, ces deux approches souffrent

de l'absence d'utilisation des connaissances. Afin de remédier à ces problèmes, plusieurs approches de coopérations régions-contours ont été proposées dans la littérature. Dans cette section, nous commençons par présenter un état de l'art de ces approches. Nous introduisons ensuite notre système multi-agents pour une approche sémantique coopérative régions-contours pour extraire les objets d'intérêt de l'image.

4.2 Les approches de segmentation coopératives régions-contours : état de l'art

Les approches de segmentation coopératives régions-contours ont été introduites dans la littérature comme une solution possible aux inconvénients des approches régions et contours appliquées séparément. Sebari et He se basent sur la manière d'intégrer la notion de coopération pour présenter trois familles d'approches régions-contours [Sebari et He, 2007].

Coopération séquentielle Cette coopération est une coopération séquentielle comme l'indique son nom. En effet, elle se compose de deux étapes. La première est soit une segmentation contour ou région. Ensuite, la deuxième étape est une segmentation de l'autre famille et qui utilise les résultats de la première segmentation. Le schéma de la figure 4.1 illustre le déroulement de ces approches.

Coopération des résultats La coopération des résultats permet comme l'indique son nom d'utiliser les résultats de deux segmentations séparées (i.e. l'une contour et l'autre région) afin d'améliorer les résultats obtenus. L'utilisation des résultats se fait en combinant les deux ou en cherchant un consensus entre les deux. L'architecture de cette approche est illustrée dans la figure 4.2

Coopération mutuelle Pour la coopération mutuelle, les algorithmes de segmentation sont exécutés en parallèle tout en s'échangeant les informations. Ces informations interchangeables permettent d'améliorer le résultat final. La figure 4.3 illustre le fonctionnement de ces approches.

La coopération régions-contours a prouvé l'apport qu'elle pouvait apporter pour la

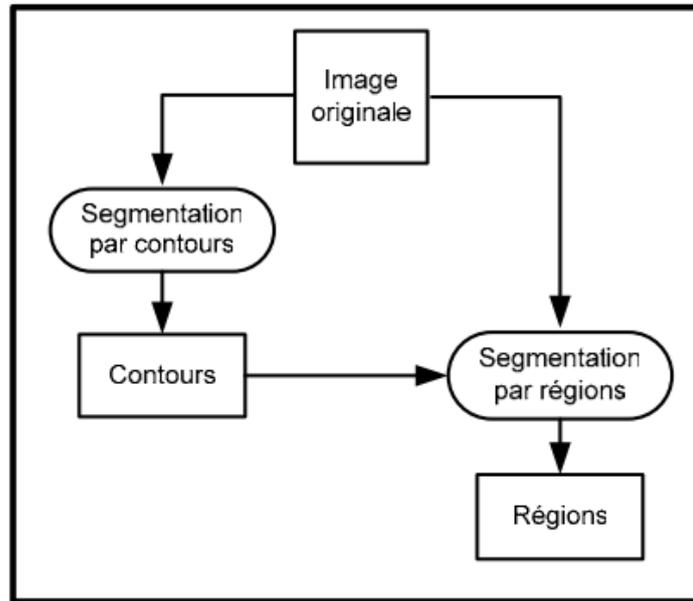


Figure 4.1 — Segmentation contours-régions séquentielle.

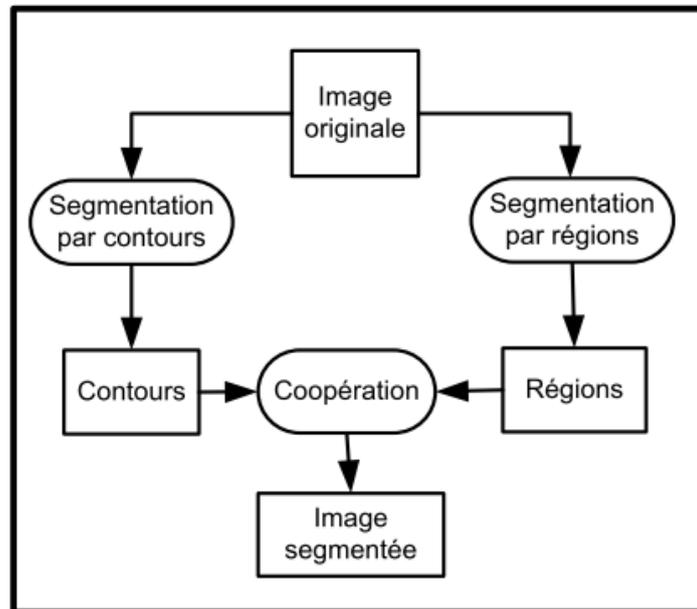


Figure 4.2 — Segmentation contours-régions par coopération de résultats.

segmentation d'images. Nous nous proposons d'intégrer une approche de coopération sémantique pour l'extraction des objets d'intérêt de l'image. Or les systèmes les plus propices pour la coopération sont les systèmes multi-agents (SMA). Dans la section suivante, nous présentons un aperçu des SMA pour ensuite introduire notre SMA sémantique coopératif régions-contours.

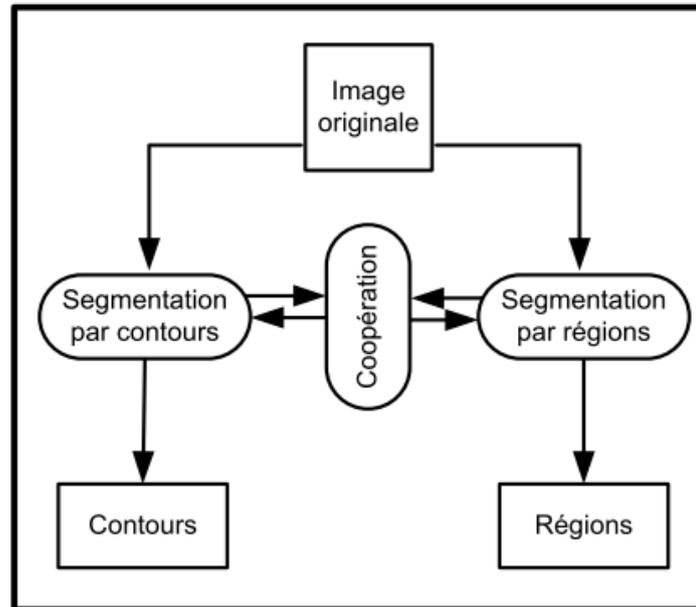


Figure 4.3 — Segmentation mutuelle contours-régions.

4.3 Les systèmes multi-agents

Les SMA ont été introduits dans le cadre de l'intelligence artificielle distribuée. Ferber et Gasser [Ferber et Gasser, 1991] définissent un SMA comme étant un système dans lequel des agents artificiels opèrent collectivement et de façon décentralisée pour accomplir une tâche.

4.3.1 Les agents

4.3.1.1 Définition

Le concept d'agent a fait l'objet de plusieurs études pendant les trois dernières décennies. Plusieurs définitions ont été proposées dans la littérature. D'après Ferber [Ferber, 1995], un agent est une entité autonome, réelle ou abstraite, qui est capable d'agir sur elle-même et sur son environnement, qui dans un univers multi-agents, peut communiquer avec d'autres agents, et dont le comportement est la conséquence de ses observations, de ses connaissances et des interactions avec les autres agents.

4.3.1.2 Types d'agents

En se fondant sur les critères d'autonomie, coopération et adaptation des agents, deux grandes familles d'agents peuvent se distinguer : les agents réactifs et les agents cognitifs.

Agents réactifs : Ce type d'agent choisit ses actions en se basant uniquement sur le percept courant. Nous parlons ici d'un système "non intelligent" d'agents. Les agents sont simples et ne possèdent pas une représentation de leur environnement, ni de mémoire, ce qui les prive d'apprentissage et de toutes anticipations aux événements. Ils sont caractérisés par l'absence de structures organisationnelles initiales prédéfinies, c'est ce qui explique leur appellation vu qu'ils agissent naturellement au moment où l'action est nécessaire. Leur comportement est de type "incitation-réponse" [Ferber, 1995]. La figure 4.4 illustre le fonctionnement des agents réactifs.

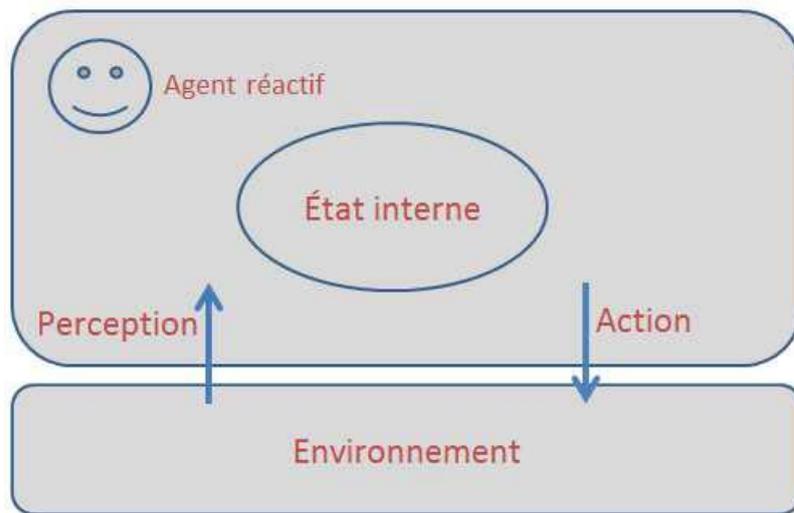


Figure 4.4 — Fonctionnement des agents réactif.

Agents cognitifs : Nous parlons ici de système d'agents "intelligents". Les agents cognitifs sont plus évolués puisqu'ils possèdent une représentation globale de leur environnement et des agents avec lesquels ils communiquent. Ils se basent aussi sur l'expérience vu qu'ils tiennent compte de leurs actions antérieures. Chaque agent possède une base de connaissances comprenant l'ensemble des informations nécessaires à l'accomplissement de sa tâche, ainsi qu'à l'interaction avec l'environnement et les autres agents [Ferber, 1995]. La figure 4.5 illustre le fonctionnement des agents cognitifs.

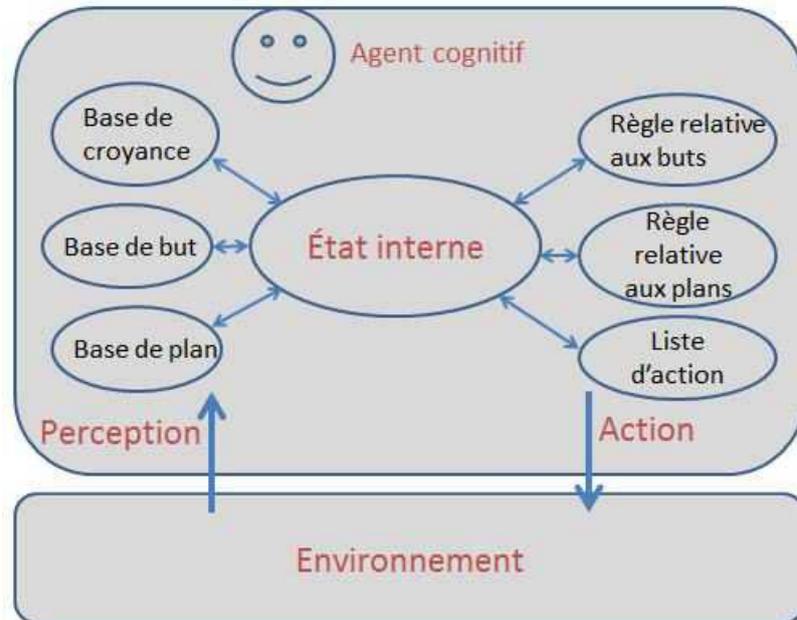


Figure 4.5 — Fonctionnement des agents cognitifs.

4.3.2 Interaction au sein des systèmes multi-agents

L'interaction est la mise en relation dynamique de deux ou plusieurs agents dans le même système. Cette interaction peut être :

- directe par envoi de messages entre les agents, elle nécessite un langage spécifique pour la communication,
- indirecte basée sur l'information entre les agents et réalisée par dépôt d'information.

Comme illustré dans la figure 4.8, qu'elle soit directe ou indirecte, l'interaction entre les différents agents au sein du même système peut être classée en trois familles : coopération, coordination et négociation.

La coopération :

Ferber [Ferber, 1995] définit la coopération comme suit : "On dira que plusieurs agents coopèrent, ou encore qu'ils sont dans une situation de coopération, si l'une des deux conditions est vérifiée :

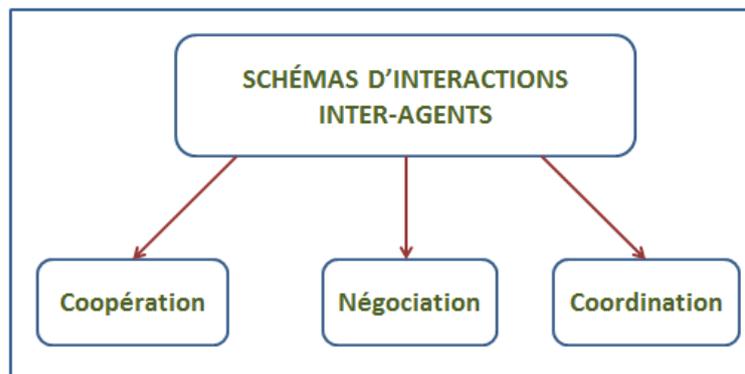
1. l'ajout d'un nouvel agent permet d'accroître les performances du groupe.
2. L'action des agents sert à éviter ou à résoudre des conflits potentiels ou actuels.

La coopération est donc indispensable afin de guider le système vers le résultat souhaité.

La coordination :

La coordination consiste à trouver une solution pour un problème donné de façon à utiliser au maximum les interactions bénéfiques et minimiser les interactions nuisibles. La coordination a aussi comme but d'organiser les activités des agents, elle peut prendre deux formes :

- Coordination centralisée : elle alloue la tâche d'exécution du plan du système à un seul agent qui joue le rôle de gestionnaire en donnant des tâches à exécuter et récupère les accusés ou les résultats de l'action.
- Coordination distribuée : elle consiste à partager les responsabilités entre les différentes entités du système, chaque agent exécute une partie du plan et coopère avec les autres agents par l'intermédiaire des messages.



La négociation : *Figure 4.6* — Interaction inter-agents.

Lors de l'interaction entre les agents, il peut y avoir des conflits. Le mécanisme intuitif pour la résolution des conflits et la coordination est inspirée du modèle humain ; c'est la négociation [Jennings *et al.*, 2001]. Il existe plusieurs types de négociation comme :

- la négociation aux enchères ;
- la négociation pour l'allocation des tâches ;
- la négociation heuristique ;
- la négociation par argumentation.

4.3.3 Les systèmes multi-agents dans l'analyse d'images

L'utilisation des SMA est récurrente dans l'analyse d'images. En effet, les auteurs utilisent les avantages et les possibilités offertes par ces systèmes afin d'améliorer les

résultats d'analyse. Plusieurs travaux ont été proposés dans la dernière décennie [Melkemi *et al.*, 2006, Mazouzi, 2008, Fleureau *et al.*, 2009, Chakroun *et al.*, 2011, Fekir et Benamrane, 2011, Moussa, 2011].

Melkemi *et al.* [Melkemi *et al.*, 2006] ont proposé un SMA pour la segmentation d'images basé sur les algorithmes génétiques. Le système est composé de K agents de segmentation interconnectés à un agent coordinateur formant une topologie en étoile. La population initiale est créée en appliquant l'algorithme K-means et en perturbant aléatoirement des pixels de l'image. Chaque agent de segmentation utilise l'algorithme ICM (iterated conditional modes) pour segmenter une partie de l'image initiale. Il transmet un paquet contenant l'image initiale, l'image segmentée et la fonction de fitness associée à cette image à l'agent coordinateur. Ce dernier sauvegarde la meilleure valeur de la fonction de fitness puis applique un croisement entre les différents parents sélectionnés à partir de la population. Ensuite, il effectue une mutation dans un ou plusieurs parents. Enfin, il évalue les résultats et retransmet les meilleures images aux différents agents de segmentation. Le système réeffectue cette opération pour un nombre défini d'itérations. Cette approche a été testée sur des images réelles et synthétiques. Les résultats sont encourageants mais la convergence des résultats est extrêmement liée aux nombres d'agents déployés, plus l'effectif s'amointrit plus le nombre d'itérations croît.

Mazouzi [Mazouzi, 2008] introduit un SMA pour la segmentation et la construction de scènes 3D. Le système proposé permet la mise en oeuvre d'une méthode de segmentation par reconstruction basée sur la coopération de deux processus : détection de point contour et reconstruction des formes géométriques. L'objectif du système proposé est de combiner entre la vision locale sur le voisinage du pixel et des états des agents qui visitent ce pixel. La mémoire d'un agent représente une perception plus large qui combinée avec l'information locale de l'image, permet d'avoir une vision globale et donc tend à favoriser une décision plus fiable.

Moussa [Moussa, 2011] présente un SMA basé sur des agents réactif pour la segmentation d'image IRM du cerveau. Après la phase de prétraitement, deux étapes sont mises en oeuvre : une segmentation approximative et une segmentation améliorative. Le système est composé de deux types d'agents :

- Agent de contrôle global : positionne les agents de contrôle locaux dans l'image en leur donnant des territoires adjacents.

- Agent de contrôle : crée des agents dédiés aux différents tissus du cerveau.

La coopération entre les agents apparaît dans le système sous trois niveaux : coopération intégrative, coopération augmentative et coopération conflictuelle. Cependant, il en résulte des motifs complexes dont la gestion est épineuse et coûteuse.

Pour résumer, l'apport des SMA est indéniable pour le traitement d'images. Son aspect distribué et sa capacité à faciliter la coopération et la communication entre les agents fait qu'ils présentent une solution plausible pour les systèmes d'analyse d'images.

4.4 Système multi-agents collaboratif régions-contours

Les approches de segmentation coopératives régions-contours sont, comme leur nom l'indique, des approches qui se basent sur la collaboration entre les deux segmentations région et contour. La phase de collaboration peut s'étendre sur tous le processus de segmentation et des interactions entre la détection de région et de contour peut permettre une meilleure extraction des objets d'intérêt. Partant du fait que les SMA présentant une architecture propice à la collaboration ainsi qu'à la communication et l'échange d'informations, nous avons choisi d'introduire un SMA pour la segmentation coopérative régions-contours. En effet, l'analyse d'images par des SMA hérite des avantages de l'approche agent, notamment la résolution d'un problème d'une manière collective et distribuée. Dans de tels systèmes, chaque agent, même s'il arrive à résoudre partiellement la tâche qui lui est alloué, flanche dans le résultat et ceci est dû au manque de compétence générale ou au manque de données globales, l'empêchant de résoudre la tâche dans son intégralité. C'est le problème des approches régions et contours qui, séparément, ne produisent pas les résultats attendus à cause des inconvénients précédemment cités. C'est pourquoi les agents sont amenés à interagir et à coopérer afin de combler leur insuffisance personnelle, en termes de compétence et de perception. Un autre apport des SMA est l'aspect du double traitement local et global qui leur permet d'avoir une double vision sur les tâches allouées. L'évolution de la dynamique du système, qui résulte de l'interaction des agents entre eux et de leur action sur leur environnement, permet de faire émerger une solution collective au problème. Et la vision locale de l'agent lui permet de gérer les problématiques locales liées à sa propre tâche.

4.4.1 Architecture du système multi-agents

Le principe de notre SMA est d'utiliser les propriétés des SMA pour catalyser la coopération entre les approches de segmentation régions et contours. Il fonctionne en deux étapes. Une première étape de prétraitement qui aura pour but d'extraire les points de départ. Cette étape sera gérée par un agent superviseur. La seconde étape est l'extraction des objets d'intérêt à partir de ces points. Un agent contour et un agent région seront affectés à chaque point de départ et vont coopérer afin d'extraire l'objet d'intérêt. La figure 4.7 illustre le fonctionnement du SMA.

4.4.2 Prétraitement

L'étape de prétraitement est une étape de préparation pour la phase d'extraction des objets. Cette étape permet d'éliminer les pixels non pertinents de l'image en se basant sur la nature de l'objet traité. Ensuite, la deuxième phase consiste en l'extraction des points de départ pour la phase d'extraction.

4.4.3 Extraction des objets.

L'étape d'extraction des objets est basée sur la coopération régions-contours. Comme nous l'avons présenté dans la section 4.2, il existe trois familles de coopération régions-contours. Dans notre système, nous utilisons une double coopération séquentielle. En effet, la croissance de région va utiliser les premiers résultats obtenus par l'agent contour comme limite pour la croissance. Ensuite, l'agent contour va être réactivé après l'étape de croissance afin de terminer le contour de l'objet en se basant sur le résultat obtenu par l'agent région. La figure 4.8 illustre le processus de coopération entre les deux agents région et contour.

4.4.4 Les agents composant le système multi-agents

Le SMA que nous proposons comporte trois agents principaux qui sont indispensables à son comportement. Ces trois agents sont :

- l'agent superviseur ;
- l'agent contour ;
- l'agent région.

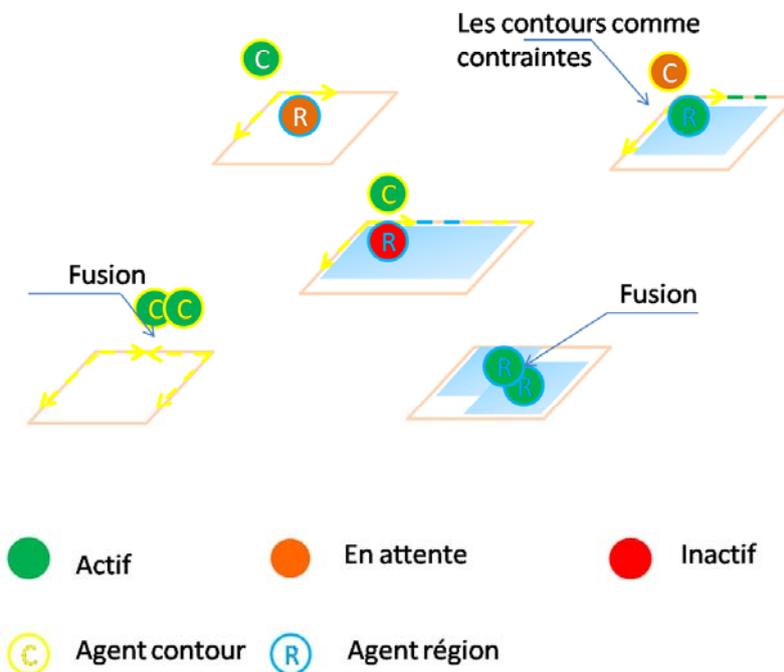
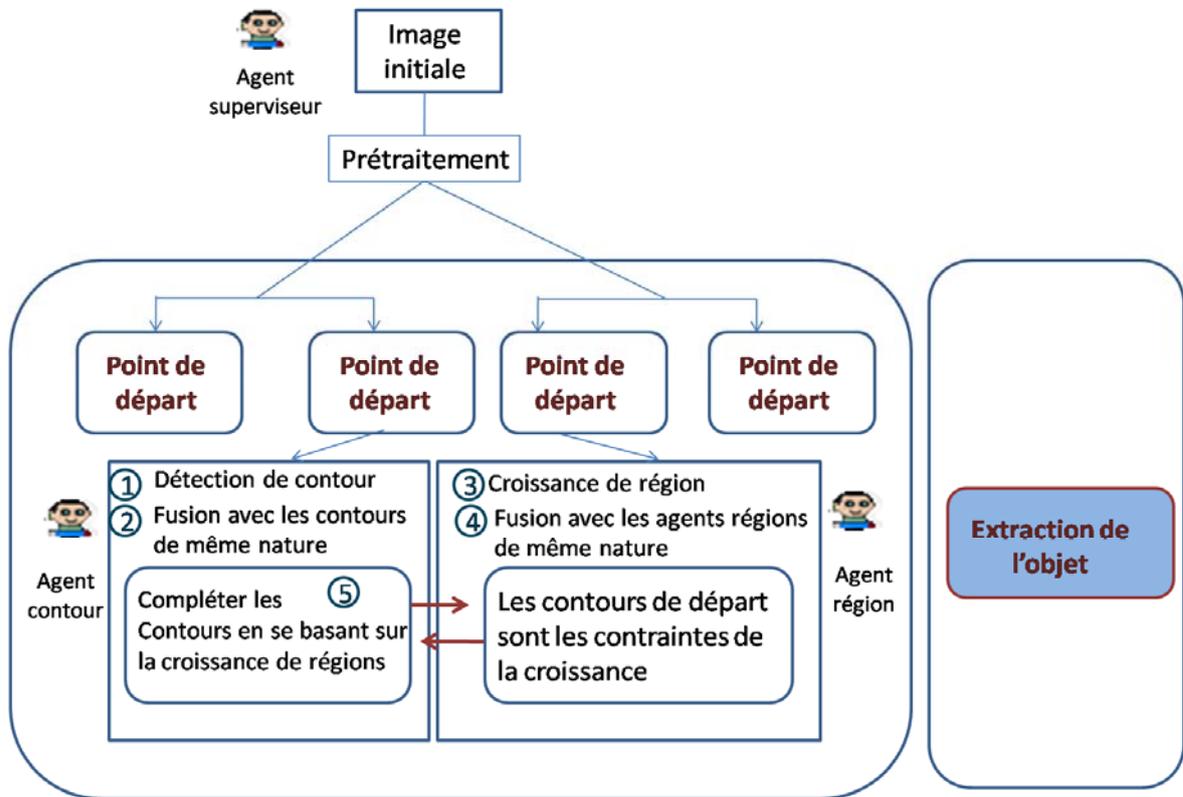


Figure 4.7 — Architecture du système multi-agents proposé.

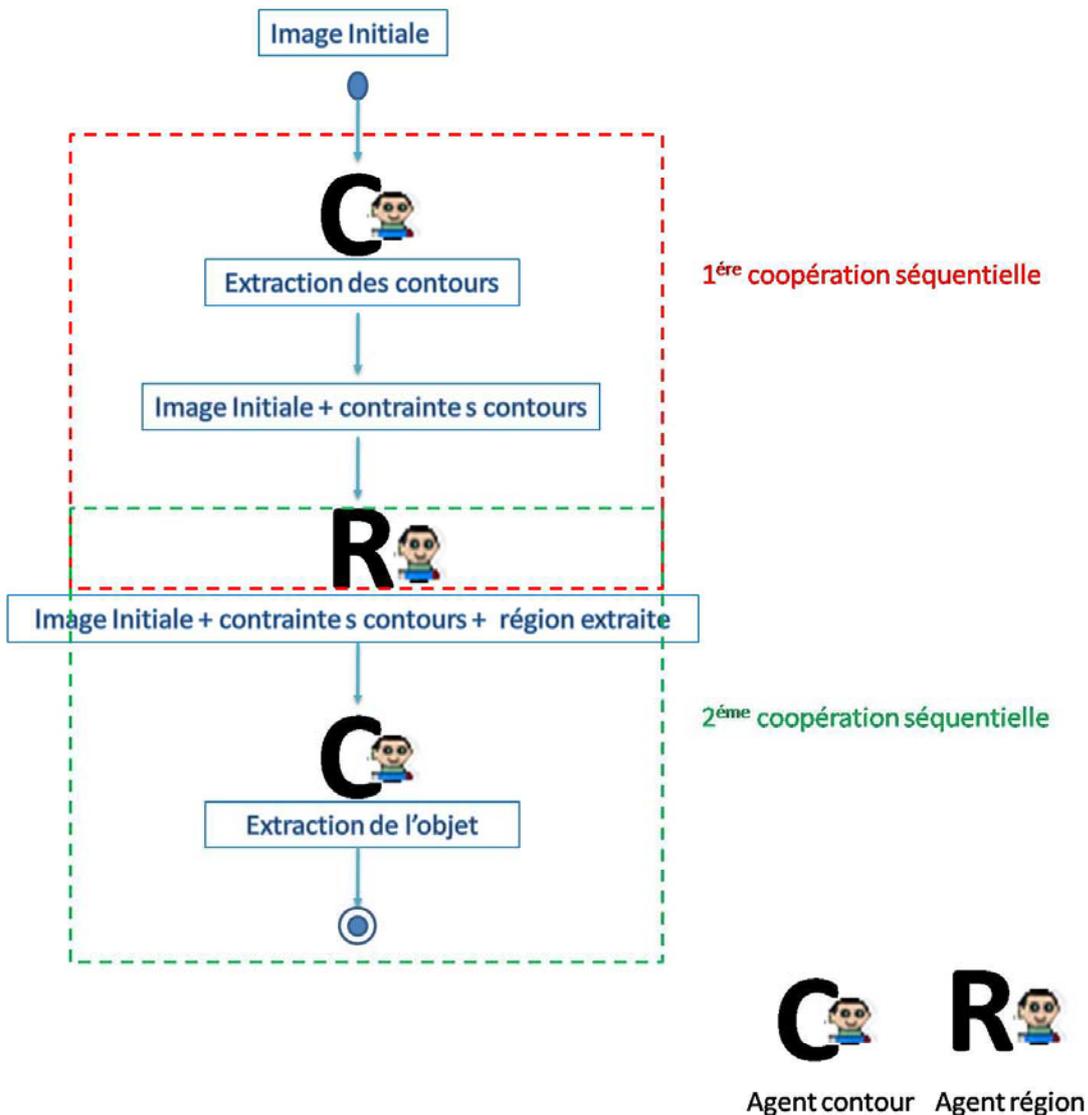


Figure 4.8 — Coopération séquentielle.

4.4.4.1 L'agent superviseur

L'agent superviseur gère la phase de prétraitement et a un rôle de coordination entre les agents contour et région. Il gère aussi l'ensemble des états des agents.

4.4.4.2 L'agent contour

L'agent contour a un double rôle visant l'extraction de l'objet final en coopérant avec l'agent région. Les deux rôles assignés à l'agent contour sont :

- L'extraction du contour initial : cette tâche peut se faire selon n'importe quel

algorithme de détection de contour. Des connaissances expertes sur l'objet à extraire peuvent être intégrées dans le processus de détection.

- La fermeture des contours en se basant sur les contours déjà extraits et sur la région détectée par l'agent région.

Le premier rôle de cet agent est donc d'extraire les contours pertinents à partir du point de départ. Soit $AC(\alpha_1, \alpha_2, \dots, \alpha_i)$ la fonction de détection de contour avec $\{\alpha_i\}_i$ l'ensemble des paramètres de cette fonction. Le résultat de la phase d'extraction est dépendant de ces paramètres.

4.4.4.3 L'agent région

Le rôle de l'agent région est de détecter l'objet tout en se limitant aux contours extraits par l'agent contour. Tous les algorithmes de croissance de régions peuvent être utilisés à condition d'y intégrer les contraintes contours. Soit $AR(\beta_1, \beta_2, \dots, \beta_j)$ la fonction de détection de régions avec $\{\beta_j\}_j$ l'ensemble des paramètres de cette fonction. Le résultat de la phase de croissance est dépendant de ces paramètres.

4.4.5 Critique du système multi-agents

Le SMA que nous avons présenté permet une double collaboration séquentielle entre un algorithme de segmentation région et un algorithme contour. Le SMA prend en entrée l'ensemble des paramètres des deux algorithmes $(\alpha_1, \alpha_2, \dots, \alpha_i, \beta_1, \beta_2, \dots, \beta_j)$. Le résultat de l'extraction reste dépendant de ces paramètres. Un autre problème majeur est l'intégration même de ces paramètres qui doivent être introduites par l'utilisateur. De plus, le bruit dans l'image ainsi que les propriétés propres à l'objet font que les paramètres qui doivent être utilisés peuvent varier d'un objet à l'autre de la même classe. Afin de palier à ces inconvénients, nous proposons l'utilisation d'un algorithme évolutionnaire local permettant d'automatiser l'intégration des paramètres de l'algorithme pour chaque objet.

4.5 Système multi-agents auto-adaptatif basé sur les algorithmes génétiques

4.5.1 Les algorithmes génétiques

Les algorithmes génétiques [Holland, 1962, Ayadi *et al.*, 2010, Sibil *et al.*, 2010, Do *et al.*, 2011] font partie de la famille des algorithmes évolutionnaires inspirés de la théorie darwinienne de la sélection naturelle. Il ne s'agit pas de trouver une solution analytique exacte mais de trouver une solution satisfaisante dans un temps raisonnable.

Appliquer un algorithme génétique consiste à faire évoluer une population initiale d'individus par des opérateurs de croisement et de mutation dans le but d'optimiser une fonction coût appelée aussi une fonction fitness.

L'entrée de l'algorithme est l'ensemble des solutions potentielles au problème. Cet ensemble est appelé "population" d'individus ou de "chromosomes". Autrement dit, le chromosome est une représentation d'un problème donné. Une première génération est choisie soit aléatoirement, soit à l'aide d'une heuristique. À partir de chaque génération, un nombre défini d'individu est choisi par un opérateur de sélection. Ces individus sont combinés deux à deux pour créer de nouveaux individus par application d'un opérateur de combinaison. La diversification des générations est assurée par l'application d'un opérateur de mutation qui garantit l'apparition de nouveaux gènes (i.e. caractères) d'une génération à une autre et évite à l'algorithme de tomber dans les optimums locaux. La figure 5.2 illustre le fonctionnement d'un algorithme génétique.

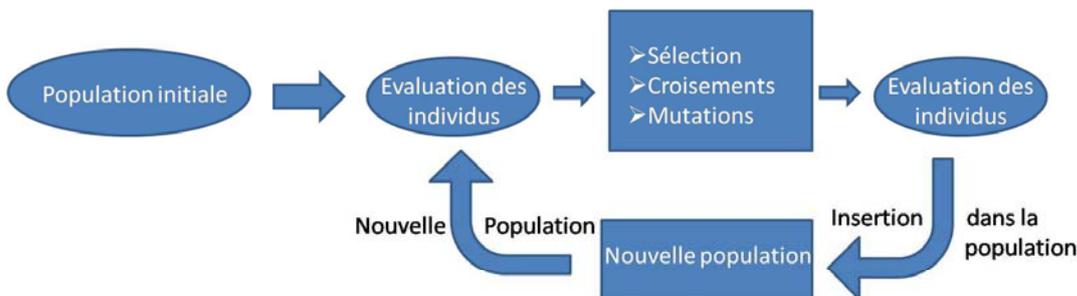


Figure 4.9 — Algorithme génétique.

4.5.1.1 Génération de la population initiale

Cette étape consiste à créer une population d'individus où chaque individu présente une solution potentiel au problème [Mendes *et al.*, 2009]. Lorsqu'il n'y a aucun à priori sur la population proche de l'optimum, la génération se fait aléatoirement en appliquant des tirages uniformes dans chacun des données associés aux composants de l'espace de recherche. Le problème principal de cette étape est le choix de la taille de la population. Si la taille est grande, le temps de calcul augmente et demande un espace mémoire important. Par contre, si la taille est petite, il y a le risque de tomber dans le problème de la convergence prématurée. Il faut donc trouver un bon compromis.

4.5.1.2 Principes de sélection

La sélection permet d'identifier les individus susceptibles d'être croisés dans une population. Il existe plusieurs techniques : sélection par rang, sélection par roulette, sélection aléatoire et sélection par tournoi. La sélection se base sur la fonction fitness (coût). La sélection est donc un processus dans lequel des individus sont choisis suivant la valeur de leur fonction de fitness qui est considérée comme la mesure de la qualité d'un individu. Plus la valeur de la fonction de fitness est élevée, plus l'individu a des chances d'être sélectionné pour la reproduction. Nous présentons quelques opérateurs pour la selection des individus.[Sivaraj et Ravichandran, 2011]

La roulette :

La sélection des individus par le système de la roulette est inspirée des roues de la loterie. En effet, à chaque individu est associé un secteur de la roue. En tournant la roue, un individu est sélectionné. La probabilité de tirer un individu est donc pondérée par sa qualité, ce qui maximise la chance des meilleurs individus d'être sélectionnés et de participer à l'amélioration de la population.

La sélection par rang :

La méthode de la sélection par rang est une variante de la méthode de sélection par la roulette. La différence réside dans la création des secteurs puisqu'elle ne repose plus sur la qualité de l'individu mais sur son rang dans la population. La méthode consiste à trier les individus en se basant sur leur qualité. Un rang est ensuite attribué à chaque individu du plus faible au plus élevé. Ensuite, chaque individu sera représenté dans la roue en fonction de son rang.



Figure 4.10 — Schéma d'une roulette

La sélection par tournoi :

Le principe de cette méthode consiste à organiser un tournoi entre les individus, les individus sont pris au hasard et le vainqueur de ce tournoi est l'individu ayant la meilleure qualité. Le point fort de cette approche est qu'elle garantit la diversification en donnant une chance au plus faible individu.

L'élitisme :

Cette méthode de sélection permet de mettre en avant les meilleurs individus de la population. Ce sont donc les individus les plus prometteurs qui vont participer à l'amélioration de la population. Cette méthode a l'avantage de permettre une convergence plus rapide des solutions. Cependant, la diversité des individus est menacée à cause du risque d'écarter des individus de piètre qualité, mais qui aurait pu apporter avec leur gène une meilleure solution.

4.5.1.3 Opérateur de croisement

Le croisement permet d'enrichir la population en manipulant les composantes des chromosomes. Le croisement se fait entre deux parents et génère un ou deux enfants avec une probabilité P_{crois} que nous appelons probabilité de croisement. Plusieurs opérateurs de croisement ont été introduits dans la littérature [Beasley *et al.*, 1993]

nous en citons :

- Croisement en 1 point (figure 4.11) : ça consiste à diviser chaque parent en deux parties à la même position, Les fragments obtenus sont alors échangés pour créer les génotypes des enfants.

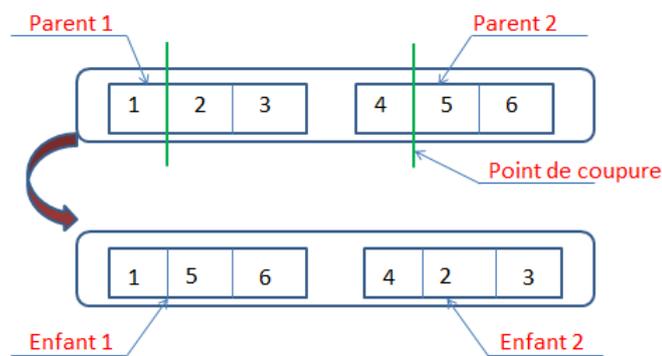


Figure 4.11 — Croisement en 1 point

- Croisement en 2 points (figure 4.12) : C'est une modification du croisement en un point. Il permet d'améliorer la flexibilité de diversification des gènes, on choisit aléatoirement deux points afin d'effectuer un échange. Les chromosomes sont coupés en deux endroits et les sous-parties obtenues sont échangées entre elles.

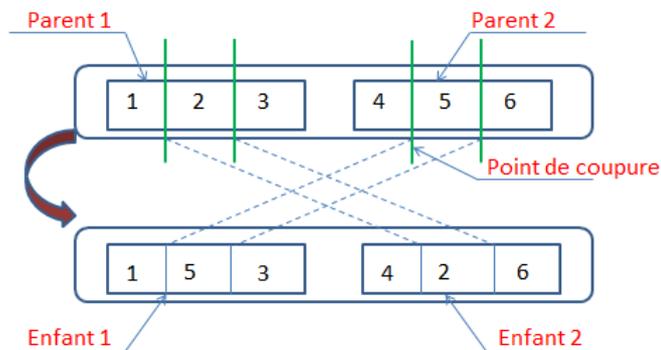


Figure 4.12 — Croisement en 2 points

4.5.1.4 Opérateur de mutation

La mutation est une modification aléatoire de la valeur d'un gène qui se produit avec une probabilité fixée de P_{mut} [Girgis *et al.*, 2009]. La figure 4.13 illustre l'opération de mutation. Cet opérateur est très intéressant dans le mécanisme de l'algorithme génétique puisqu'il apporte l'aléa nécessaire à une exploration efficace de l'espace et permet de quitter les extrema locaux.

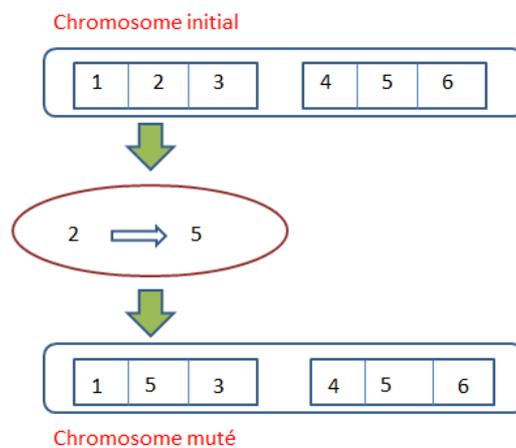


Figure 4.13 — Principe de l'opérateur de mutation

4.5.1.5 Fonction de fitness

La fonction de fitness est un pion essentiel dans le processus d'optimisation [Nelson *et al.*, 2009]. Cette fonction est dépendante du problème à résoudre et du codage choisi pour les chromosomes. Pour chaque chromosome, elle attribue une valeur numérique proportionnelle à la qualité de l'individu en tant que solution.

4.5.1.6 Condition d'arrêt

Afin d'aboutir à un résultat satisfaisant tout en prenant en compte le coût de l'algorithme génétique, il faut bien gérer la condition d'arrêt de ce dernier en faisant un compromis entre le résultat obtenu et le coût. Pour l'arrêt de l'algorithme, deux conditions sont prises en compte :

- La première condition concerne le nombre de générations, notre algorithme s'arrête une fois qu'il atteint nb générations.

- La deuxième condition d'arrêt concerne la variation de la qualité entre un nombre de générations consécutives.

4.5.2 Architecture du système multi-agents auto-adaptatif

Le problème rencontré par notre SMA réside dans sa dépendance aux paramètres des deux agents contour et région. Afin de résoudre ce problème, nous appliquons localement l'algorithme génétique dans notre SMA. La figure 4.14 illustre le schéma global de l'intégration de l'algorithme génétique dans notre SMA.

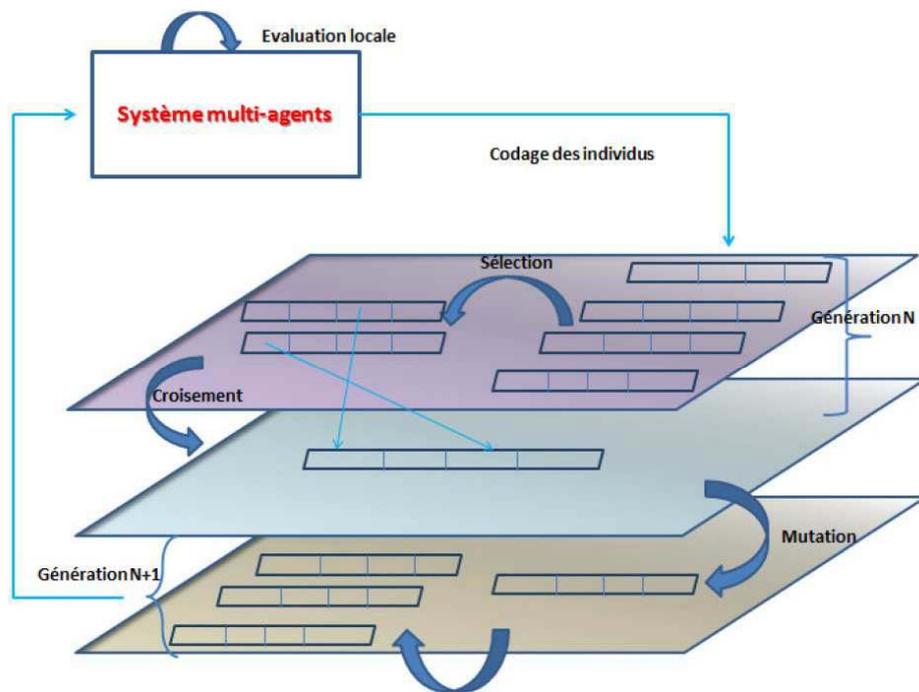


Figure 4.14 — Schéma global de notre système multi-agents intégrant l'algorithme génétique.

L'intégration de l'algorithme génétique à notre SMA est gérée au niveau de l'agent superviseur. Au moment du lancement des deux agents région et contour, l'agent superviseur crée une population initiale d'individus spécifiques. Chaque individu représente une solution potentielle au problème : il comporte n gènes qui représentent les valeurs respectives $(\alpha_1, \alpha_2, \dots, \alpha_i, \beta_1, \beta_2, \dots, \beta_j)$. L'évaluation de l'individu se fait en évaluant la région détectée par la coopération des deux agents. Les différentes étapes de l'algorithme génétique (croisement, mutation, sélection) seront appliquées par l'agent superviseur sur les individus de chaque génération jusqu'à atteindre un nombre de génération

spécifique au domaine d'application. À ce stade, le meilleur individu est retenu et l'objet est détecté avec les paramètres optimaux. Cette procédure a l'avantage de pouvoir déterminer pour chaque objet le jeu de paramètre qui lui convient. Sachant que même en cherchant à détecter le même objet dans l'image, ces objets risquent de varier sur différentes propriétés telles que la forme, la taille ou les propriétés radiométriques. L'application locale de l'algorithme génétique peut aussi gérer les problèmes dus au bruit.

4.5.3 Apport de l'application de l'algorithme génétique dans l'amélioration de notre système multi-agents

Après avoir introduit notre algorithme génétique et présenté les différentes étapes qui le composent, nous résumons les avantages de l'utilisation de l'algorithme génétique dans notre SMA.

- *un SMA coopératif et auto-adaptatif* : l'utilisation de l'algorithme génétique dans le SMA a permis l'automatisation des choix des paramètres.

- *choix local des paramètres pour chaque objet* : notre SMA est capable de déterminer pour chaque objet de l'image une combinaison spécifique des paramètres des deux algorithmes région et contour permettant une bonne détection. En effet, les objets peuvent varier. Ce traitement local à l'objet permet une spécification du traitement et une meilleure adaptation des agents par rapport aux bruits et aux changements que peut subir l'objet

4.6 Conclusion

Dans ce chapitre, nous avons présenté un SMA permettant la collaboration entre deux agents contour et région. Cette collaboration est gérée par un agent superviseur. Le premier constat que nous avons eu sur le SMA était sa dépendance aux différents seuils qui représentaient les entrées des agents contours et régions. Le choix de ces seuils n'était pas automatique et dépendait de l'utilisateur. Le jeu de seuils introduit était utilisé pour tous les objets. Cependant, les différentes variations que peut subir l'objet, que ce soit à cause du bruit ou des propriétés propres à l'objet, fait que l'utilisation

d'un jeu de seuils unique ne donnait pas entière satisfaction. Afin d'y remédier, nous avons intégré un algorithme génétique local pour chaque objet. Cet aspect local qui caractérise les agents a permis un choix propre pour chaque objet. Dans le chapitre suivant, nous adapterons notre système pour la détection des bâtiments en y intégrant les connaissances expertes extraites à partir du dictionnaire de données.

Un système multi-agents évolutionnaire auto-adaptatif pour l'extraction des bâtiments

5.1 Introduction

Dans ce chapitre, nous introduisons une adaptation du SMA évolutionnaire proposé dans le chapitre 4 afin d'extraire les bâtiments. En effet, les étapes du SMA proposé sont spécifiques à l'objet d'intérêt. Nous intégrons dans le SMA les propriétés spécifiques aux bâtiments afin d'améliorer l'extraction. Dans ce cadre, nous commençons par présenter l'architecture du SMA adapté aux bâtiments. Ensuite, nous identifions les différents composants de l'algorithme génétique.

5.2 Un système multi-agents pour l'extraction des bâtiments

Le SMA proposé dans le chapitre 4 est composé de deux parties. Une première partie de prétraitement et une deuxième partie d'extraction des objets. Afin d'adapter

notre système à l'extraction des bâtiments, nous gardons la même architecture tout en introduisant des propriétés spécifiques aux bâtiments dans les différents agents du SMA. Nous introduisons aussi un nouvel agent dans l'étape de prétraitement : c'est l'agent NDVI. Cet agent va permettre d'éliminer un ensemble de pixels non pertinents et d'aider l'agent superviseur dans sa tâche. La figure 5.1 illustre l'adaptation du SMA pour les bâtiments.

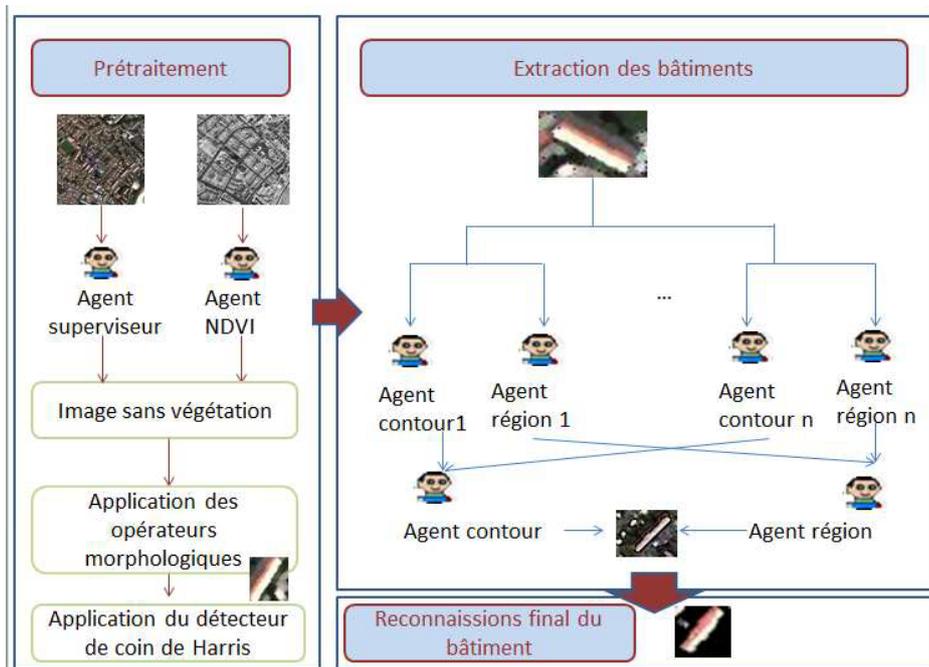


Figure 5.1 — Architecture du SMA pour l'extraction des bâtiments.

Quatre agents forment le SMA. Les fonctionnalités de ces agents sont les suivantes :

- **L'agent superviseur** qui gère le prétraitement et la coopération entre les différents agents.
- **L'agent NDVI** qui élimine les pixels non pertinents de type végétation
- **L'agent contour** qui possède deux rôles :
 1. Détection des contours : les pixels appartenant aux contours des bâtiments sont détectés en utilisant des propriétés spécifiques à l'agent. Le contour extrait est ensuite envoyé à l'agent région et jouera le rôle de contraintes pour la croissance.
 2. La correction de contour : l'agent va compléter le contour manquant en utilisant le contour déjà extrait et la région extraite par l'agent région.

- **L'agent région** qui exécute un algorithme d'extraction de région en se fondant sur les contraintes contours fournies par l'agent contour.

5.2.1 Prétraitement

Comme nous l'avons déjà présenté, l'étape de prétraitement est une phase fondamentale dans notre système d'extraction de bâtiment. Elle permet de préparer le terrain pour les agents contour et région. La figure 5.2 illustre les trois phases de prétraitement de l'image.

Considérant la variété du contenu des images satellites, il est très utile dans l'étape de prétraitement d'éliminer les pixels non pertinents. Connaissant les propriétés des végétations qui présentent une valeur de l'indice différentiel normalisé de végétation (NDVI) élevée par rapport aux autres classes, l'agent superviseur déploie un agent NDVI capable de déterminer pour chaque pixel de l'image son NDVI [Singh *et al.*, 2012]. Le NDVI est défini comme suit :

$$NDVI = \frac{(PIR - Rouge)}{(PIR + Rouge)} \quad (5.1)$$

avec PIR et Rouge représentant respectivement la mesure de radiométrie de la bande proche infra-rouge et de la bande rouge. L'ensemble des pixels détectés par cet agent sera transmis à l'agent superviseur. Ce dernier éliminera ces pixels de la liste des pixels de l'image afin d'éviter les traitements inutiles par le reste des agents dans les phases ultérieures. Ces pixels seront aussi considérés comme étant des contraintes et seront utilisés dans la phase d'extraction.

Une fois les pixels de végétation éliminés, l'agent superviseur applique à l'image l'opérateur morphologique de dilatation [Ghandour, 2010]. Le but de cette procédure est d'améliorer le contraste des objets de l'image et spécialement les toits des bâtiments et de corriger les contours irréguliers et discontinus. Pour appliquer la dilatation d'une image couleur, les couleurs sont considérées comme des labels associés à chaque pixel. Ensuite ces valeurs sont utilisées afin d'établir un ordre total dans l'espace des couleurs. Pour ce travail, nous avons appliqué la méthode la plus utilisée qui est la méthode de réduction d'ordre [Comer et Delp, 1999].

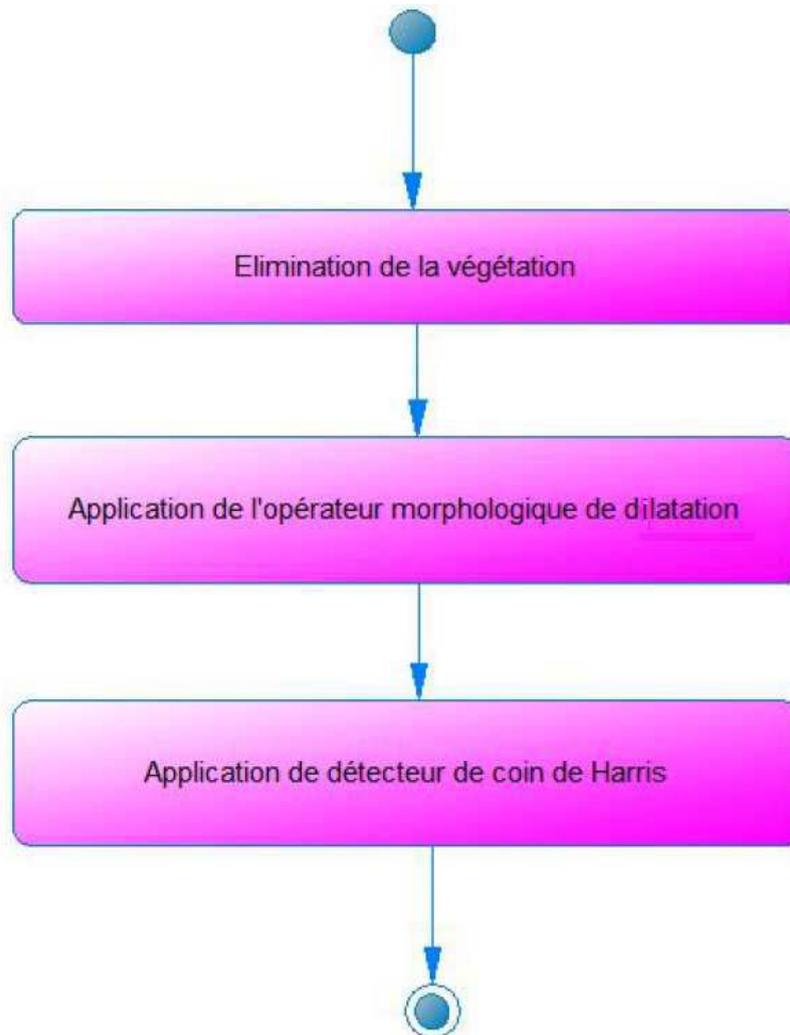


Figure 5.2 — Phase de prétraitement pour l'extraction des bâtiments.

La dernière étape dans la phase de prétraitement consiste à détecter les points de départ des agents contour et région. Sachant que les bâtiments présentent majoritairement des formes rectangulaires ou carrées qui sont caractérisées par des coins, nous nous sommes orientés vers le détecteur de coin de Harris [Harris et Stephens, 1988, Pei *et al.*, 2010, Bangalore *et al.*, 2011, Kang *et al.*, 2011]. Ce détecteur est connu par sa résistance aux bruits et son invariance à la rotation, translation et à la variation au niveau du contraste. L'agent superviseur applique ce détecteur sur l'ensemble des pixels de l'image, privée des pixels de végétation afin d'extraire les candidats susceptibles d'être traités comme étant des coins de bâtiments.

5.2.2 Extraction des bâtiments

La deuxième étape est celle de l'extraction des bâtiments. À partir de l'ensemble des coins de bâtiments potentiels extraits dans la phase de prétraitement, l'agent superviseur lance deux agents : un agent contour et un agent région. Le processus de traitement:

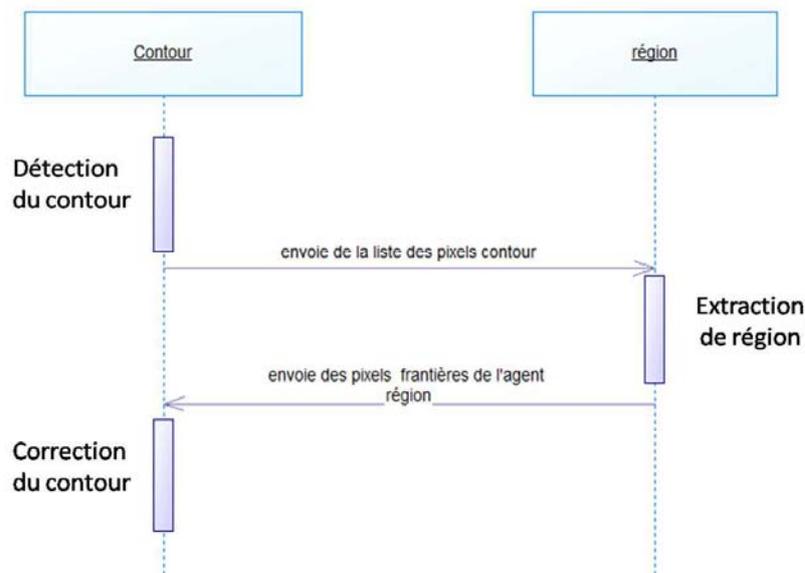


Figure 5.3 — Coopération entre un agent région et un agent contour pour l'extraction d'un bâtiment.

L'agent contour commence le processus d'extraction en cherchant les pixels constituant les contours des bâtiments. Le traitement de cet agent est inspiré du détecteur de contour de Canny [Canny, 1986, Medina-Carnicer *et al.*, 2011]. En effet, l'agent commence par appliquer deux étapes du détecteur de Canny à savoir : un filtre gaussien et une estimation du gradient. Ensuite, l'agent cherche à partir du pixel courant les pixels non encore traités dans une fenêtre 3×3 tout en favorisant les pixels qui sont sur la direction du gradient. Les pixels retenus sont les pixels qui appartiennent à un contour de bâtiment. Ces pixels doivent donc valider deux hypothèses :

- Hypothèse 1 : cette hypothèse cherche à s'assurer que le pixel en question appartient à un contour. Donc, si un pixel est susceptible d'appartenir à un contour, il ne doit pas parcourir une région homogène. Afin de valider ce critère, nous calculons la distance des deux fenêtres à droite et à gauche du pixel en question en suivant la direction du gradient. La distance de Bhattacharya est utilisée pour garantir un minimum

de distance entre les deux fenêtres à gauche et à droite du pixel courant. La distance de Bhattacharya est définie comme suit :

$$\frac{1}{4(m_1 - m_2)^t(\Gamma_1 + \Gamma_2)^{-1}(m_1 - m_2)} + \frac{1}{2} \ln \frac{\frac{1}{2}|\Gamma_1 + \Gamma_2|}{\sqrt{|\Gamma_1||\Gamma_2|}} \quad (5.2)$$

où m_i et Γ_i sont les moyennes et les covariances des distributions. Plusieurs travaux ont comparé la distance de Bhattacharya avec d'autres distances. Ces travaux ont montré que cette distance propose les meilleurs résultats [Reyes-Aldasoro et Bhalerao, 2006]. La distance seuil permettant de valider cette hypothèse est notée S_{min} . La figure 5.4 illustre ce processus entre les deux régions A et B.

- Hypothèse 2 : la deuxième hypothèse vérifie que le point en question est le contour d'un bâtiment et non celui d'un autre objet. Pour ce faire, la distance de Bhattacharya entre une des deux fenêtres à gauche ou à droite du pixel courant et une fenêtre extraite par l'expert et représentant un bâtiment doit être inférieure à une distance maximale notée S_{max} afin d'assurer que l'objet soit bien un bâtiment.

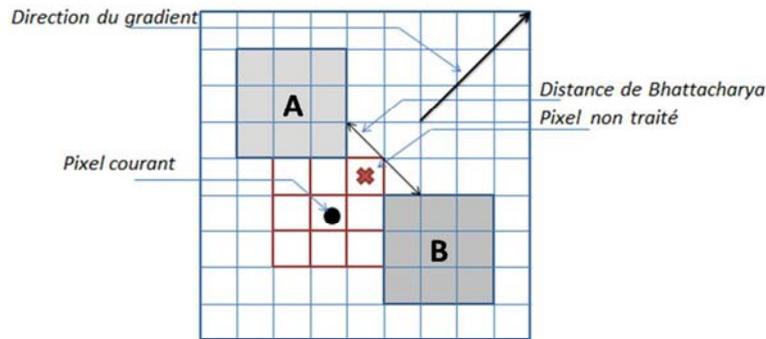


Figure 5.4 — La distance de Bhattacharya.

L'agent contour continue la procédure de détection jusqu'à ce qu'il soit dans l'un de ces deux cas :

- lorsqu'il croise un autre agent contour, alors les deux agents fusionnent et donnent naissance à un nouvel agent qui contient la liste des pixels détectés par les agents parents.
- lorsque l'agent contour n'a plus de pixel à traiter.

Une fois son premier rôle terminé, il donne la main à l'agent région. Le rôle principal de l'agent région est de déterminer l'ensemble des pixels qui constituent le corps du

bâtiment. Pour ce faire, l'agent applique la procédure de croissance de région par la technique du meilleur ajustement [Lang *et al.*, 2009]. Cet algorithme consiste à faire évoluer une région en ajoutant de nouveaux pixels vérifiant une variation de l'homogénéité inférieure à un seuil S_{hom} défini. La croissance de région s'arrête une fois que l'agent région rencontre un autre agent région, dans ce cas, les deux agents fusionnent pour donner un nouvel agent région qui admet les pixels résultant de la croissance de ces parents. Le deuxième critère d'arrêt est l'absence de pixel validant le critère de croissance.

Dès que l'agent contour reçoit les pixels extraits par l'agent région, il continue la détection du contour, la procédure s'arrête une fois que le contour est fermé et le bâtiment est détecté.

Il faut mentionner que chacun des deux agents oriente et pose des contraintes sur le fonctionnement de l'autre agent : l'agent contour doit au cours de sa détection tenir compte des pixels déterminés par l'agent région, généralement les contours des bâtiments ne se trouvent pas dans une région homogène mais entre deux régions hétérogènes. L'agent région à son tour utilise les contours fixés par l'agent contour pour arrêter sa croissance puisque les contours des objets se trouvent à la limite entre deux régions.

Notre système est maintenant en place. Les trois paramètres dont il dépend sont S_{min} , S_{max} et S_{hom} . Comme nous l'avons présenté dans le chapitre précédent, cette dépendance est préjudiciable pour le SMA vu qu'elle rend le système non automatique. Un deuxième point évoqué est la variabilité au sein même des objets qui fait que les paramètres qui sont convenables pour un objet i peuvent ne plus l'être pour un objet j même si ces deux objets appartiennent à la même classe. Pour les cas des bâtiments, cette règle est toujours maintenue. Pour y remédier, la solution que nous avons proposé est d'intégrer localement un algorithme génétique permettant d'optimiser le traitement pour chaque bâtiment.

5.3 Algorithme évolutionnaire local automatisant le SMA pour l'extraction des bâtiments

L'intégration de l'algorithme évolutionnaire à notre SMA va permettre l'automatisation et le traitement local des agents. Cette intégration est gérée par l'agent superviseur. Lors du lancement des deux agents région et contour, l'agent superviseur crée une population initiale d'individus. Chaque individu représente une solution potentielle au paramétrage des deux agents région et contour. Un individu comporte 3 gènes qui représentent les valeurs respectives de S_{min} , S_{max} et S_{hom} . Une fois la population initiale créée, l'agent superviseur lance pour chaque individu un agent contour et un agent région à partir du point de Harris avec les paramètres contenus dans les gènes de l'individu (S_{min} , S_{max}) pour l'agent contour et S_{hom} pour l'agent région. L'évaluation de chaque individu se fait en évaluant le résultat de l'extraction du bâtiment sur la base de ses caractéristiques. Les différentes étapes de l'algorithme génétique (i.e. croisement, mutation, et sélection) seront appliquées par l'agent superviseur sur les individus de chaque génération jusqu'à atteindre un nombre de génération égale à N pour l'algorithme génétique. A ce stade, le meilleur individu est retenu et le bâtiment est détecté avec les paramètres de ce dernier. Cette procédure a l'avantage de pouvoir déterminer pour chaque bâtiment le jeu de paramètre qui lui convient, sachant que les bâtiments dans une image varient de point de vue radiométrie et contraste. La figure 5.5 illustre les différentes étapes du SMA évolutionnaire que nous présentons par la suite.

5.3.1 Codage des individus

Soit i un individu (i.e. qui représente un génotype dans l'algorithme génétique) de la population. Pour coder chaque individu, nous avons choisi le codage réel vu la facilité d'application des opérateurs génétiques de combinaison et de mutation. De plus, la nature des paramètres à optimiser facilite leur codage sous forme de gène réel. Chaque individu i est codé avec un vecteur composé de trois paramètres qui sont les gènes S_{min} , S_{max} et S_{hom} .

5.3.2 Fonction de fitness

Le choix de la fonction fitness a un grand impact sur la convergence de l'algorithme génétique. Notre choix a été guidé par la nature même de notre algorithme. En effet,

cette fonction va évaluer le résultat obtenu à partir de l'ensemble des paramètres affectés aux deux agents région et contour. Nous évaluons donc le résultat de l'extraction d'un bâtiment. Or, comme nous l'avons déjà présenté dans la section 3.6, l'une des propriétés fondamentales des bâtiments est qu'ils ont généralement une forme bien définie. Cette forme est généralement rectangulaire ou carrée.

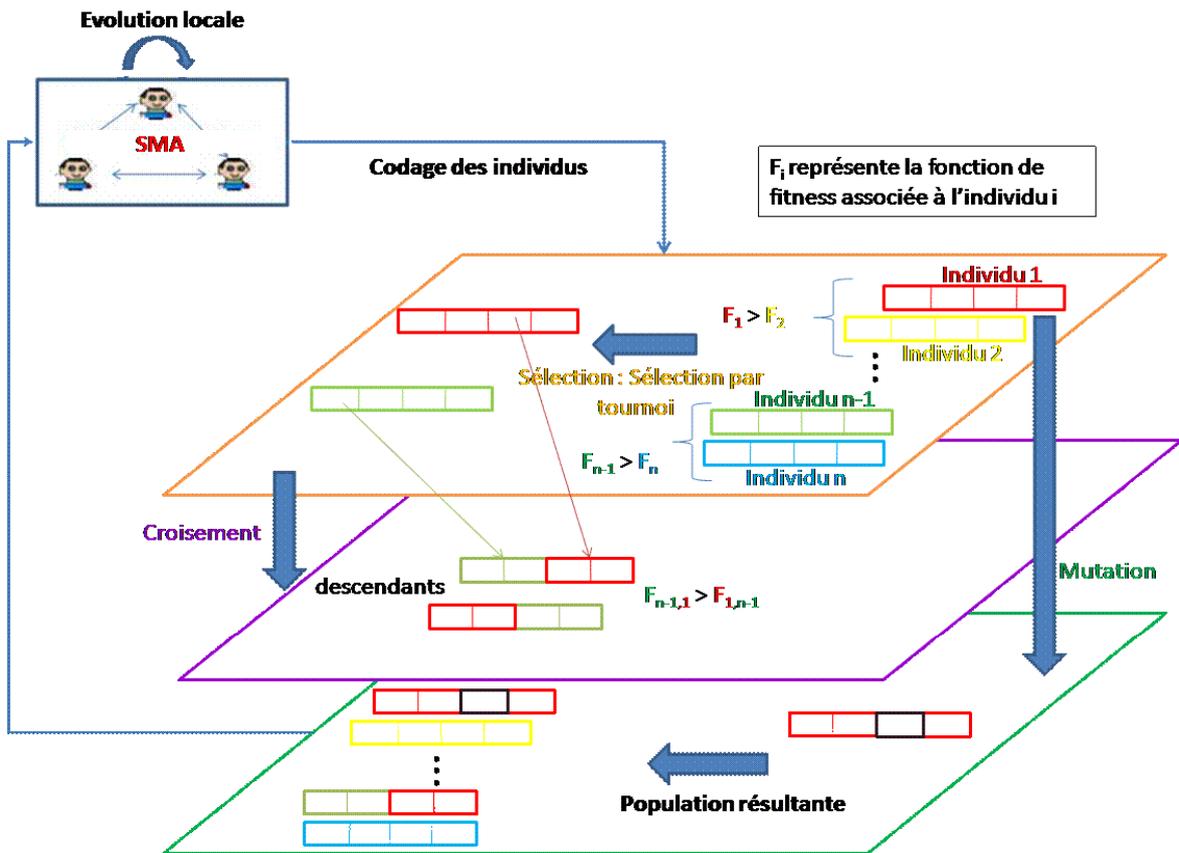


Figure 5.5 — SMA Évolutionnaire local pour l'extraction des bâtiments.

Les figures 3.3 et 3.1 montrent deux représentations de la classe "Bâtiment" à savoir les pavillons et les bâtiments de l'habitat collectifs. Le dictionnaire des données décrit les bâtiments d'habitat collectifs comme étant "des bâtiments de forme rectangulaire (étroit et allongé au sol) ou de forme carré de volume allongé vers le haut". Les bâtiments aussi ont une superficie au sol variant entre une valeur minimale et maximale. En se basant sur toutes ces connaissances, nous avons conçu la fonction de fitness suivante :

$$F(i) = \text{Aire} \times \left(\frac{\text{Rec}}{1 + \text{Hom}} \right) \quad (5.3)$$

cette fonction met en relation trois paramètres : la rectangularité Rec , la superficie $Aire$ et l'homogénéité Hom . Notons r le bâtiment extrait par l'individu i . Nous rappelons la définition de $Aire$:

$$Aire = \begin{cases} 1 & \text{si } aire(r) \in [min_{aire}, max_{aire}] \\ \frac{aire(r)}{min_{aire}} & \text{if } aire(r) < min_{aire} \\ \frac{max_{aire}}{aire(r)} & \text{if } aire(r) > max_{aire} \end{cases} \quad (5.4)$$

Cette fonction permet de catalyser la croissance si la taille de l'objet qui est entrain de croître est inférieure à la borne inférieure. En effet, dans l'intervalle $[0, min_{aire}[$ la fonction est croissante. Si l'aire de la région est dans l'intervalle $[min_{aire}, max_{aire}]$, la fonction est alors constante et elle reste neutre vu que sa valeur se stabilise à 1 permettant ainsi aux autres attributs d'être plus discriminant. Lorsque cette valeur dépasse la borne supérieure (i.e. dans l'intervalle $]max_{aire}, +\infty[$), cette fonction devient décroissante et elle pénalise la croissance.

Rec est la mesure de rectangularité de la région déterminée par les deux agents. Dans ce travail, nous avons eu recours à l'algorithme de Rotating Calipers [Toussaint, 1983, Chang *et al.*, 2011] qui permet de mesurer l'indice de rectangularité d'une région qui varie entre $[0,1]$.

Hom désigne l'indice d'homogénéité de la région r , calculé par la valeur de l'écart type de la région. Une valeur de l'écart type faible indique que les pixels de région sont assez proches en terme de radiométrie et une valeur élevée indique que la région n'est pas tout à fait homogène.

La fonction fitness a tendance à favoriser les individus qui ont une mesure de rectangularité élevée dont la superficie varie entre $[min_{aire}, max_{aire}]$, et une mesure d'homogénéité assez élevée caractérisée par un faible écart type.

5.3.3 Opérateurs génétiques utilisés

5.3.3.1 L'opérateur de sélection

Comme présenté dans la section précédente, l'opérateur de sélection permet de sélectionner les individus afin d'améliorer la population. L'idée de base de cet opérateur est de promouvoir les meilleurs individus en leur permettant de transmettre leur gènes aux futures générations. Chaque individu est évalué en se basant sur sa fonc-

tion de fitness. Plusieurs opérateurs de sélection ont été introduits dans la littérature [Sivaraj et Ravichandran, 2011].

Notre choix s'est orienté vers la sélection par tournoi binaire. Cette méthode a l'avantage de promouvoir la diversité tout en gardant les meilleurs individus. En effet, à chaque fois deux individus sont sélectionnés aléatoirement. Cette sélection aléatoire permet d'assurer la diversité de la population vu que même les individus ayant une faible fonction de fitness peuvent être sélectionnés. Le choix du meilleur individu parmi les deux choisis permet de conserver la qualité des individus.

5.3.3.2 L'opérateur de croisement

Le croisement a pour but d'enrichir la diversité de la population en manipulant la structure des chromosomes. En effet, l'information génétique des deux parents est fusionnée pour créer de nouveaux individus. Les croisements sont envisagés avec deux parents et génèrent deux enfants. Nous avons choisi ici un croisement à 1 seul point. Le choix de cet opérateur est dû à la petite taille de notre chromosome (individu composé de 3 paramètres).

5.3.3.3 L'opérateur de mutation

L'opérateur de mutation est une déformation aléatoire du gène d'un individu permettant de préserver la diversité de la population et d'éviter de tomber dans un optimum local [Mendes *et al.*, 2009]. La mutation est caractérisée par une probabilité informant sur la fréquence à laquelle les gènes d'un chromosome sont mutés. Le choix de cette probabilité tend généralement vers une valeur minimale afin de ne pas trop influencer sur la diversité de la population et de laisser le processus de croisement s'en occuper. Nous définissons cette valeur à $P_{mut} = 0.03$.

5.3.4 Fonctionnement de l'algorithme génétique

L'approche génétique que nous proposons est intégrée au niveau du SMA. Dans cette section, nous décrivons les différentes étapes qui composent notre approche.

5.3.4.1 Création de la population initiale

Pour la création de la population initiale, l'agent superviseur crée la population initiale. Chaque individu présente une solution de notre problème (i.e. trouver la combinaison de paramètre optimale pour les deux agents contours et région qui permet d'extraire avec précision les bâtiments à partir de l'image). Ils comportent 3 gènes S_{min}, S_{max} et S_{hom} , dont les valeurs varient entre $[0, 255]$. L'espace de recherche est $E = [0, 255]^3$.

5.3.4.2 Évaluation de la population

Pour chaque individu, un agent contour et un agent région sont lancés à partir du point de Harris et détectent l'objet bâtiment avec les paramètres compris dans les gènes de cet individu. Le résultat de la détection est retenu pour une région. Cette région sera évaluée par la fonction de fitness F que nous avons introduit pour déterminer la qualité de cet individu. Cette procédure est répétée jusqu'à épuisement de tous les individus de la population.

5.3.4.3 Évolution des populations

À partir d'une population N déjà évaluée, deux opérateurs génétiques peuvent être appliqués pour passer de la génération N à une autre génération $N + 1$. Un opérateur de croisement, qui consiste à tirer deux individus aléatoirement, combiner les deux individus père pour obtenir deux nouveaux individus fils qui sont évalués et le meilleur d'entre eux remplacera le moins bon des parents (i.e. par rapport à la fonction de fitness). Une fois l'opérateur de combinaison appliqué, un autre opérateur de mutation est mis en oeuvre. Un individu est sélectionné au hasard, et l'un de ces gènes est modifié pour créer un nouvel individu avec une nouvelle valeur pour la fonction fitness.

Une fois ces deux opérateurs appliqués, nous obtenons une nouvelle population qui contient des individus plus performants que les individus de la population précédente.

5.4 Expérimentation et évaluation

Dans cette section, nous allons présenter un exemple d'exécution de notre approche sur les deux images présentées dans le chapitre 3. Nous illustrons l'évolution des étapes de notre système sur ces images. Nous présentons les deux étapes fondamentales du

système : une étape de prétraitement et une étape d'extraction des bâtiments.

L'étape de prétraitement est une étape primordiale dans notre approche, cette étape est effectuée par l'agent superviseur dans le but de préparer l'image pour l'étape d'extraction des bâtiments. Elle comporte trois phases : élimination de la végétation, dilatation de l'image puis détection des coins de Harris.

Pour l'étape d'élimination de la végétation, l'agent superviseur dispose d'une image quatre bandes (rouge , bleu , vert et proche infrarouge) comme l'illustre les figures 5.6 et 5.7, il lance l'agent NDVI pour déterminer les pixels végétation (figure 5.8) et les élimine de l'image initiale.



Figure 5.6 — Les images de test.

la deuxième procédure dans l'étape d'initialisation consiste à dilater l'image pour améliorer le contraste des objets de l'image. Pour ce faire, l'opérateur morphologique de dilatation est utilisé, la figure 5.9 montre les résultats de l'application de cet opérateur sur les deux images test .

Enfin, la dernière phase du prétraitement consiste en l'extraction des coins de Harris. Au cours de cette phase, l'agent superviseur applique le détecteur de coin de Harris pour l'extraction des coins des bâtiments, ces pixels vont servir plus tard pour le lancement des deux agents contour et région et la délimitation des bâtiments de l'image. La figure 5.10 représente les pixels de Harris (en bleu) extraits à partir de l'image.

La deuxième étape consiste en l'extraction des bâtiments en se basant sur les coins



(a) image proche infrarouge site1



(b) image proche infrarouge site2

Figure 5.7 — Images proche infrarouge des deux sites.



(a) image ndvi site1



(b) image ndvi site2

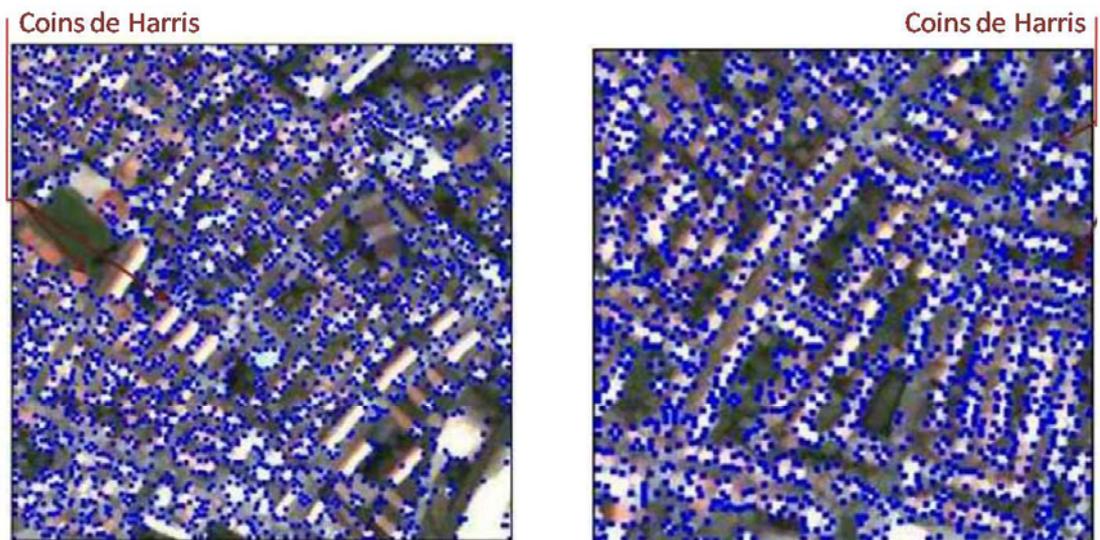
Figure 5.8 — Résultat de l'indice de végétation des images test .



(a) image dilatée site 1

(b) image dilatée site 2

Figure 5.9 — Résultat de l'opérateur de dilatation sur les images test .



(a) coins de Harris pour le site 1

(b) coins de Harris pour le site 2

Figure 5.10 — Résultat de l'application de l'opérateur de Harris sur les images test.

de Harris détectés. Cette étape est le résultat d'une coopération entre un agent région et un agent contour, nous avons déjà mentionné que notre approche d'extraction des bâtiments est basée sur l'application d'un algorithme génétique pour la détermination des paramètres qui permettent une meilleure extraction des bâtiments moyennant les différents agents du système. Le résultat de l'extraction sur les deux image 5.6 et 5.7

Tableau 5.1 — Évaluation du rappel et de la précision du SMA évolutionnaire pour la classe "Bâtiment"

	Bâtiment
Rappel	0.8967
Précision	0.8770

est illustré dans les figures 5.11(a) et 5.11(b).



(a) Site 1 : Zone éparses

(b) Site 2 : Zone denses

Figure 5.11 — Extraction de la classe "Bâtiment"

L'évaluation numérique présentée dans le tableau 5.1 introduit les résultats d'extraction de bâtiments. Les résultats obtenus montrent que l'approche est compétitive que ce soit du point de vue rappel que précision.

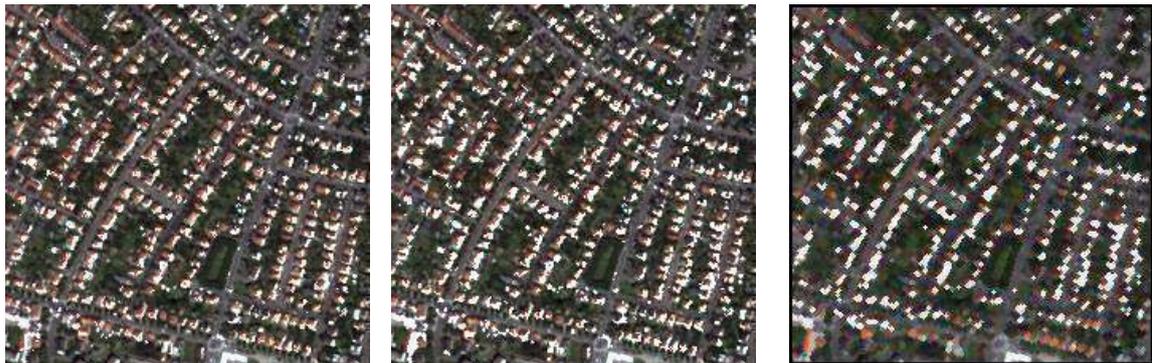
Nous illustrons maintenant l'apport de l'utilisation de l'algorithme génétique et l'évolution des résultats de génération en génération. La figure 5.12 montre l'amélioration des résultats le long de l'évolution du SMA.

L'évaluation de cette évolution à travers l'étude du rappel et de la précision (Tableau 5.1 démontre l'augmentation continue des valeurs des deux métriques.

La figure 5.13 montre la pente de croissance des deux courbes de rappel et de précision entre les 3 générations étudiées.

Tableau 5.2 — Evolution de l'algorithme génétique

Génération	Rappel	Précision
10	85	120
20	145	66
30	179	16



(a) 10^{ème} génération

(b) 20^{ème} génération

(c) 30^{ème} génération

Figure 5.12 — Extraction des bâtiments au cours de l'évolution de l'algorithme génétique.

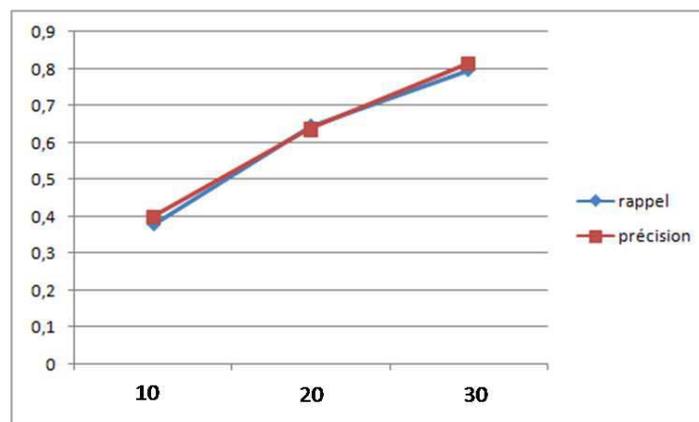


Figure 5.13 — Évolution du rappel et de la précision.

5.5 Conclusion

Dans cette section, nous avons adapté notre SMA évolutionnaire auto-adaptatif pour permettre l'extraction des bâtiments. Nous avons introduit un nouvel agent qui a profité de la nature et du contexte des zones urbaines pour éliminer des pixels non pertinents. Nous avons aussi présenté les différents opérateurs que nous avons utilisé et nous

avons argumenté le choix de ces derniers en nous fondant sur les propriétés intrinsèques des bâtiments. Dans la section suivante, nous présenterons l'expérimentation des approches proposées dans cette thèse sur des images urbaines à très haute résolution de la ville de Strasbourg.

Validation et expérimentation dans le cadre de l'analyse de scène urbaine

6.1 Introduction

Dans les précédents chapitres de cette thèse, nous avons présenté différents algorithmes permettant une interprétation objets collaborative, offrant ainsi une meilleure interaction entre les deux principales étapes des approches d'analyse d'images orientées objets, à savoir, la construction des objets et leur identification. Ces approches ont été évaluées sur deux exemples uniquement. Dans cette section, nous présentons une évaluation plus détaillée de ces approches en proposant plusieurs jeux de test et en utilisant un ensemble de métriques d'évaluation permettant de mieux évaluer ces approches.

6.2 Données utilisées

L'image sur laquelle nous avons testé notre approche est une image à très haute résolution à quatre bandes spectrales de résolution spatiale 2,44 - 2,88 m/px et une bande panchromatique de résolution 0.7m couvrant la zone urbaine de l'agglomération de Strasbourg. Cette image nous a été fournie par le Laboratoire Image, Ville, Environnement, et a été capturée par le satellite américain Quickbird (figure 6.1) en 2008.



Figure 6.1 — Vue spatiale du satellite Quickbird

QuickBird est un satellite très haute résolution d'observation de la Terre, propriété de DigitalGlobe lancé en 2001 par un lanceur Delta II à partir de la base de Vandenberg en Californie. Le tableau 6.1 illustre les caractéristiques du satellite.

Satellite	Résolution	Bande spectrale	Altitude	Inclinaison	Orbite	Fauché
Quickbird	0,61/0,72m PAN 2,44/2,88m MUL	R, G, B, PIR	450km	98 degrés	polaire	16,5km

Tableau 6.1 — Informations sur le satellite Quickbird

La figure 6.2 introduit l'image intégrale en couleur tandis que la figure 6.3 représente l'image proche infrarouge de la même zone.

Afin d'évaluer la robustesse des approches que nous avons proposé, nous avons extrait des images de plusieurs zones de la ville de Strasbourg en essayant de diversifier les images de test en nous basant sur les critères de densité (i.e. images denses et images éparées) de taille (différentes tailles de fenêtres) et de diversité des objets (plusieurs types de bâtiments et de routes avec différentes tailles et formes). Nous présenterons, dans ce qui suit, l'ensemble des images relatives à plusieurs sites de la ville de Strasbourg, à savoir, Cronenbourg ouest, la Meinau, Oberhausbergen, Eckbolsheim et Quartier des Quinze.

L'évaluation des résultats de nos approches sur les images introduites dans la section précédente ne peut être faite sans l'aide d'une vérité terrain ou d'un expert du domaine.

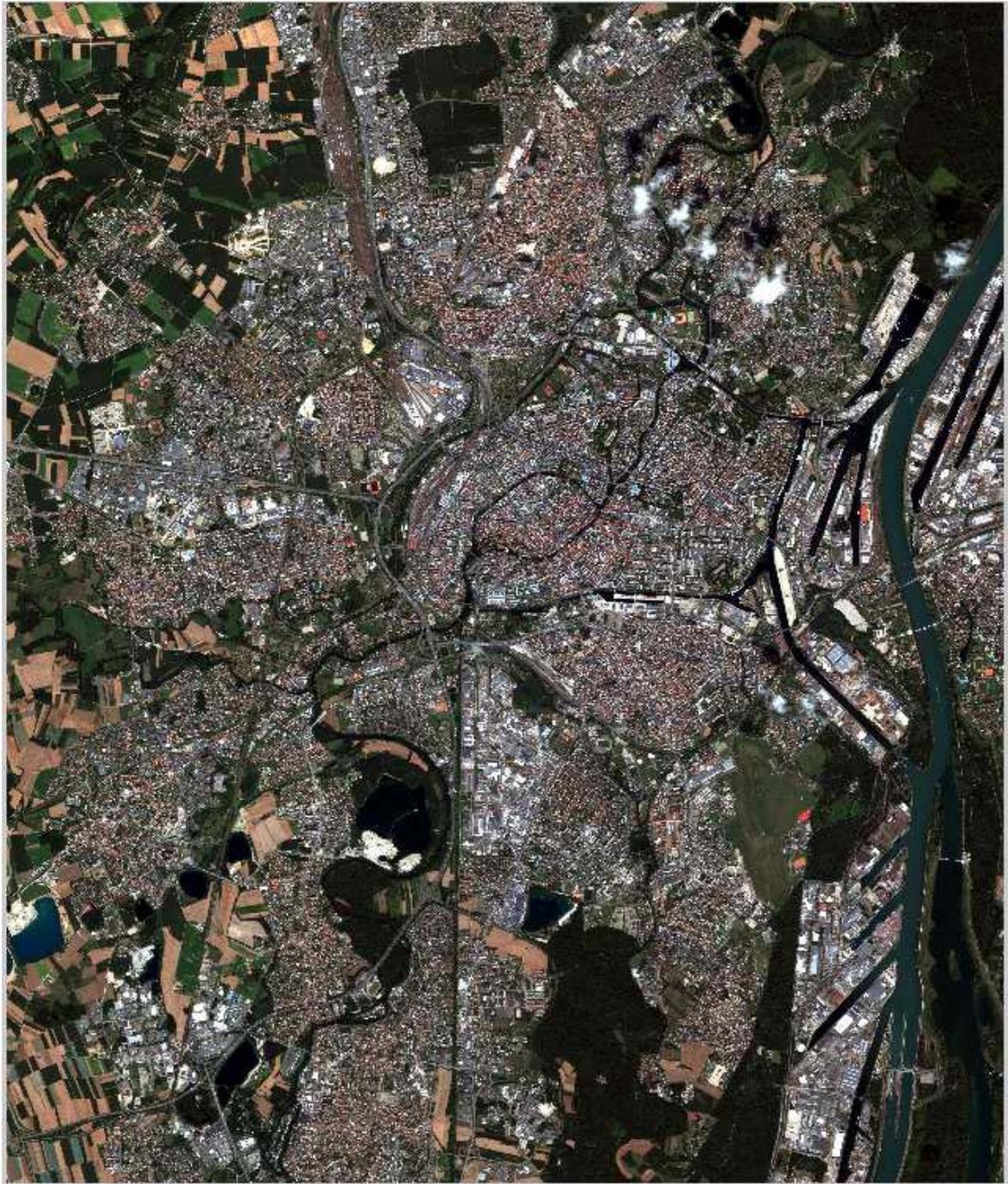


Figure 6.2 — Illustration de l'image intégrale de la zone d'étude en couleur

La vérité terrain est une représentation correcte du résultat d'une application idéale de l'algorithme de reconnaissance ou d'extraction. Afin de valider nos images, nous avons été amenés à extraire la vérité terrain correspondante à partir d'une carte numérique

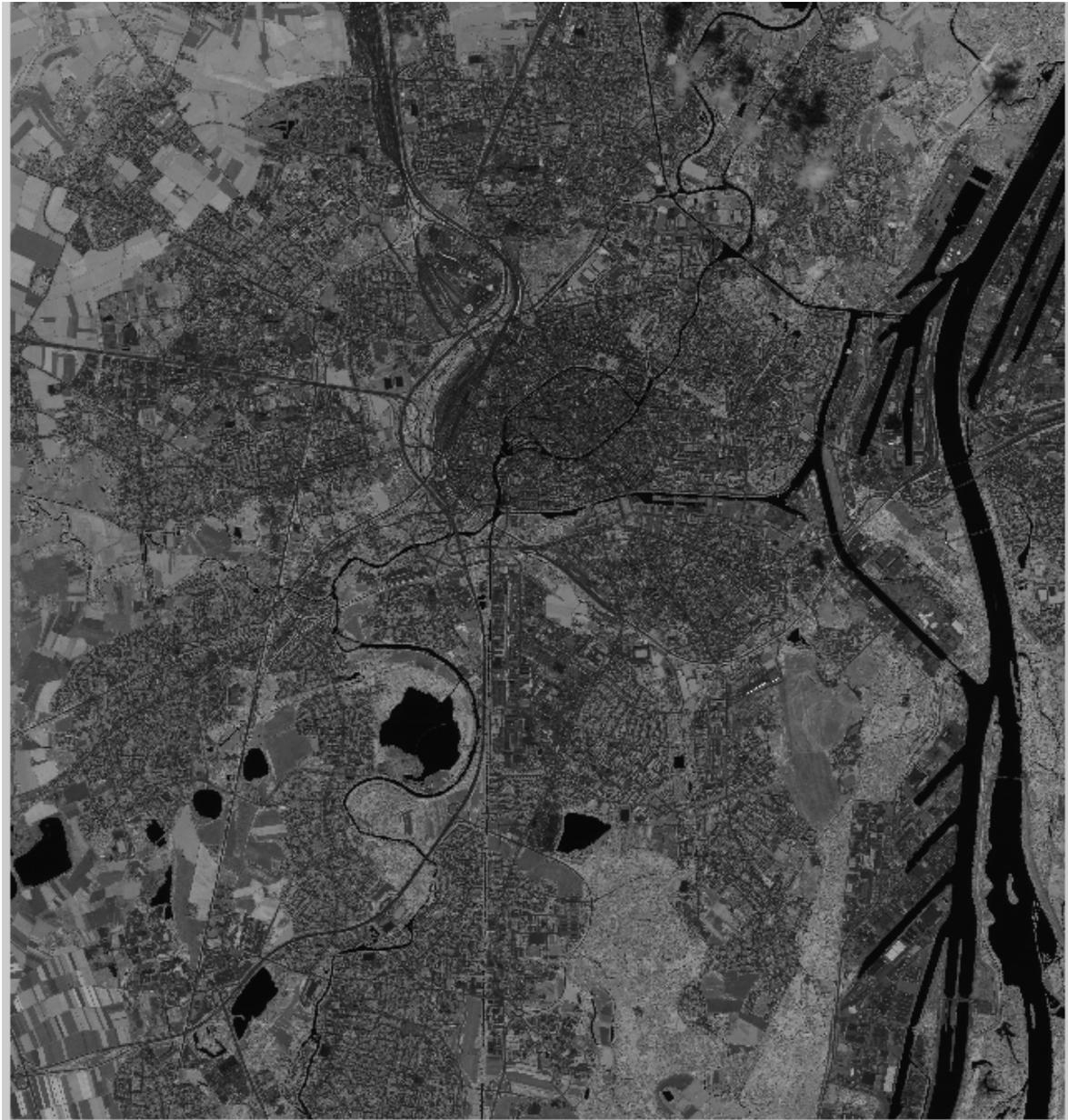


Figure 6.3 — Illustration de l'image intégrale de la zone d'étude en proche infra rouge (PIR)

de la ville de Strasbourg¹. L'ensemble des images de test ainsi que les vérités terrain associées à ces images sont introduites dans les figures suivantes.

La figure 6.4 représente un extrait de la cité de Cronembourg ouest avec la vérité terrain (figure 6.4(b)) correspondante. Ce site (figure 6.4(a)) représente une zone moyennement éparsée contenant des pavillons et des bâtiments de l'habitat collectif. Concernant

1. <http://www.carto.strasbourg.eu/>



(a) site Cronenbourg ouest

(b) Site Cronenbourg ouest vérité terrain

Figure 6.4 — Image de test de Cronenbourg ouest.

les routes, cette image est caractérisée par la présence d'un parking et d'un réseau routier non quadrangulaire assez complexe.

La figure 6.5 représente la cité de la Meinau et sa vérité terrain (figure 6.5(b)). Ce site représente une zone dense avec deux types de bâtiments : un groupement de pavillons et des bâtiments de l'habitat collectif de couleur blanche. Le réseau routier est quasiment quadrangulaire mais il représente des rayons de courbures assez amplifiés.

La figure 6.6 représente un site de la cité d'Oberhausbergen caractérisé par un réseau routier quadrangulaire et une densité élevée de bâtiments. La figure 6.6(b) illustre la vérité terrain de ce site.

Pour l'image de la cité de Quartier des Quinze (la figure 6.7(a)), nous remarquons la présence d'un mélange de pavillons et de bâtiments de l'habitat collectif de différentes couleurs, formes et tailles. Nous distinguons dans ce site des ronds-points, des parkings ainsi qu'un réseau routier très complexe. La vérité terrain correspondante est introduite par la figure 6.7(b).

La figure 6.8(a) représentant Eckbolsheim est très dense et présente une grande variété de pavillons et de bâtiments de l'habitat collectif de forme assez spécifique et déviant de la forme rectangulaire ainsi que d'un réseau routier très complexe caractérisé par une irrégularité aiguë des formes.



(a) site Quartier des Quinze

(b) site Quartier des Quinze vérité terrain

Figure 6.7 — Images de test du Bas Rhin.



(a) site Eckbolsheim

(b) site Eckbolsheim vérité terrain

Figure 6.8 — Images de test d'Eckbolsheim.

Le tableau 6.2 récapitule les caractéristiques des images de test choisies. Les critères que nous avons visés sont la densité des images, la variabilité des objets d'intérêt ainsi que la complexité.

Tableau 6.2 — Caractéristiques des images de test

Sites	Densité	Variabilité	Complexité
Cronenbourg ouest	+/-	+	+
La Meinau	-	+/-	+/-
Eckbolsheim	+	+/-	-
Quartier des Quinze	+/-	+	+
Oberhausbergen	+	+	+

Nous pouvons constater que les images de test présentent divers caractéristiques qui peuvent permettre une bonne évaluation de la robustesse des approches testées.

6.3 Métriques d'évaluation

Plusieurs métriques d'évaluation ont été proposées dans la littérature. Les métriques les plus utilisées sont le rappel et la précision que nous avons introduits dans le chapitre 3. Cependant, et même si ces deux métriques permettent d'évaluer la capacité du système à restituer l'ensemble des objets pertinents ainsi que sa capacité à ne restituer que des objets pertinents, elles manquent de précision par rapport à la qualité des résultats obtenus. Afin d'y remédier, Mayunga *et al.* [Mayunga *et al.*, 2007] ont repris les deux mesures de *rappel* et de *précision* en y intégrant une nouvelle métrique. Ces métriques ainsi proposées prennent en compte les objets partiellement détectés permettant une meilleure estimation des performances de l'approche évaluée. Dans ces travaux, les auteurs présentent une méthode pour l'analyse qualitative des résultats obtenus par un algorithme de détection de bâtiment. La fiabilité de l'approche est déterminée en calculant le pourcentage des bâtiments extraits.

Afin de calculer les métriques introduites, Mayunga *et al.* commencent par présenter quatre mesures qui sont :

- OCE qui représente le nombre d'objets correctement extraits ;
- OPE qui représente le nombre d'objets partiellement extraits ;
- ONE qui représente le nombre d'objets non extraits ;

– FA qui représente les fausses alarmes.

La métrique présentée dans [Mayunga *et al.*, 2007] est donnée par :

$$TEO = \frac{OCE}{OCE + OPE + ONE} * 100$$

La TEO évalue la complétude d'extraction des objets. Une autre métrique est définie dans le même travail qui permet une évaluation qualitative des approches d'extractions, les paramètres d'évaluations qualitatives sont définis comme suit :

$$Exactitude = \frac{OCE}{OCE + FA}$$

$$Qualité = \frac{OCE}{OCE + OPE + ONE + FA}$$

L'*Exactitude* permet d'évaluer la performance du système à bien délimiter les objets d'intérêts en estimant le nombre d'objets extraits par rapport au nombre d'objets correctement extraits et le nombre de fausses alarmes. La métrique de *Qualité* permet de combiner les aspects des deux autres mesures pour résumer la performance du système et la qualité d'extraction.

6.4 Évaluation de la classe Bâtiment

Nous présentons dans cette section les résultats d'extraction de bâtiments pour les images de test. Les figures illustrent les résultats d'extraction de l'approche sémantique hiérarchique ainsi que celles du SMA évolutionnaire pour les différents sites de test. Les résultats obtenus (figures 6.14, 6.15 et 6.11) montrent l'intérêt des deux approches proposées. En effet, et malgré les différentes propriétés qui caractérisent les images de test, les résultats restent assez compétitives et permettent une bonne détection des bâtiments.

De manière à quantifier la précision des résultats des deux approches proposées, nous avons extrait les métriques précédemment introduites. Nous avons obtenu une bonne complétude qui avoisine les 80 %. Les valeurs varient entre 85% et 90% pour la première approche et 82 % et 87 % pour la deuxième. Ceci prouve que malgré la diversité des sites choisis, les approches restent stables du point de vue complétude et permettent un bon pourcentage d'extraction.



(a) Résultat d'extraction de l'approche hiérarchique du Site de Cronenbourg ouest.



(b) Résultat d'extraction du SMA du Site de Cronenbourg ouest.



(c) Résultat d'extraction de l'approche hiérarchique du Site de La Meinau.



(d) Résultat d'extraction du SMA du Site de La Meinau.

Figure 6.9 — Résultat de l'extraction des bâtiments dans les images de test I.



(a) Résultat d'extraction de l'approche hiérarchique du Site d'Oberhausbergen.



(b) Résultat d'extraction du SMA du Site d'Oberhausbergen.



(c) Résultat d'extraction de l'approche hiérarchique du Site du Bas Rhin.



(d) Résultat d'extraction du SMA du Site du Bas Rhin.

Figure 6.10 — Résultat de l'extraction des bâtiments dans les images de test II.



(a) Résultat d'extraction de l'approche hiérarchique du Site d'Eckbolsheim. (b) Résultat d'extraction du SMA du Site d'Eckbolsheim.

Figure 6.11 — Résultat de l'extraction des bâtiments dans les images de test III.

Concernant l'exactitude, les valeurs avoisinent les 90 % pour les deux approches. Ceci montre qu'elles permettent une bonne localisation des bâtiments dans la scène d'étude et favorisent une extraction qui dévie des erreurs de confusion entre les différentes classes d'intérêts.

Finalement, la *Qualité* atteint les 80% pour la première approche et les 70 % pour la seconde. Ceci s'explique par les bâtiments partiellement détectés.

La figure 6.12 illustre la comparaison quantitative entre les résultats des deux approches proposées pour l'extraction des bâtiments. Les résultats qualitatifs montrent que l'approche hiérarchique est meilleure pour les trois critères. Ceci s'explique par le fait que l'approche évolutionnaire n'est pas une approche exacte mais qu'elle cherche à trouver un bon paramètre dans un temps raisonnable. Ceci se répercute sur la qualité des bâtiments extraits. Pour les bâtiments non extraits, l'approche SMA se base sur l'utilisation d'un détecteur de Harris. Cependant, les bâtiments dont les coins ne sont pas extraits par le détecteur de Harris ne seront pas localisés ce qui accroît le nombre de bâtiments non détectés. Cependant, l'approche SMA semble avoir une meilleure précision géométrique (i.e. les bâtiments détectés ne débordent pas) selon l'évaluation visuelle. Ceci s'explique par la nature même de cette approche qui utilise les contours comme contraintes pour la croissance.

Tableau 6.3 — Évaluation quantitative de la classe bâtiment.

Sites	OCE	OPE	ONE	FA	TEO	Exactitude	Qualité %
Approche Hiérarchique site Cronenbourg ouest	251	27	17	17	85.08	93.65	80.44
Approche SMA site Cronenbourg ouest	249	16	30	31	84.40	88.92	76.38
Approche Hiérarchique site La Meinau	367	22	20	20	89.73	94.83	85.54
Approche SMA site La Meinau	320	20	59	19	80.20	94.39	76.55
Approche Hiérarchique site Eckbolsheim	296	29	12	9	87.83	97.04	85.54
Approche SMA site Eckbolsheim	295	14	28	38	87.5	88.5	78.6
Approche Hiérarchique site Quartier des Quinze	98	7	3	10	90.74	90.74	83.05
Approche SMA site Quartier des Quinze	95	6	7	12	87.96	88.78	79.16
Approche Hiérarchique site Oberhausbergen	82	7	5	8	87.23	91.11	80.39
Approche SMA site Oberhausbergen	78	9	7	12	82.97	86.66	73.58

Afin de situer nos deux approches par rapport aux récentes approches proposées dans la littérature, nous proposons de les comparer en utilisant les deux critères de TEO et d'exactitude. Le tableau 6.4 présente les résultats des différentes approches.

La figure 6.13 permet de bien situer les résultats obtenus par nos deux approches par rapport à six autres méthodes de détection de bâtiments dans des images satellitaires

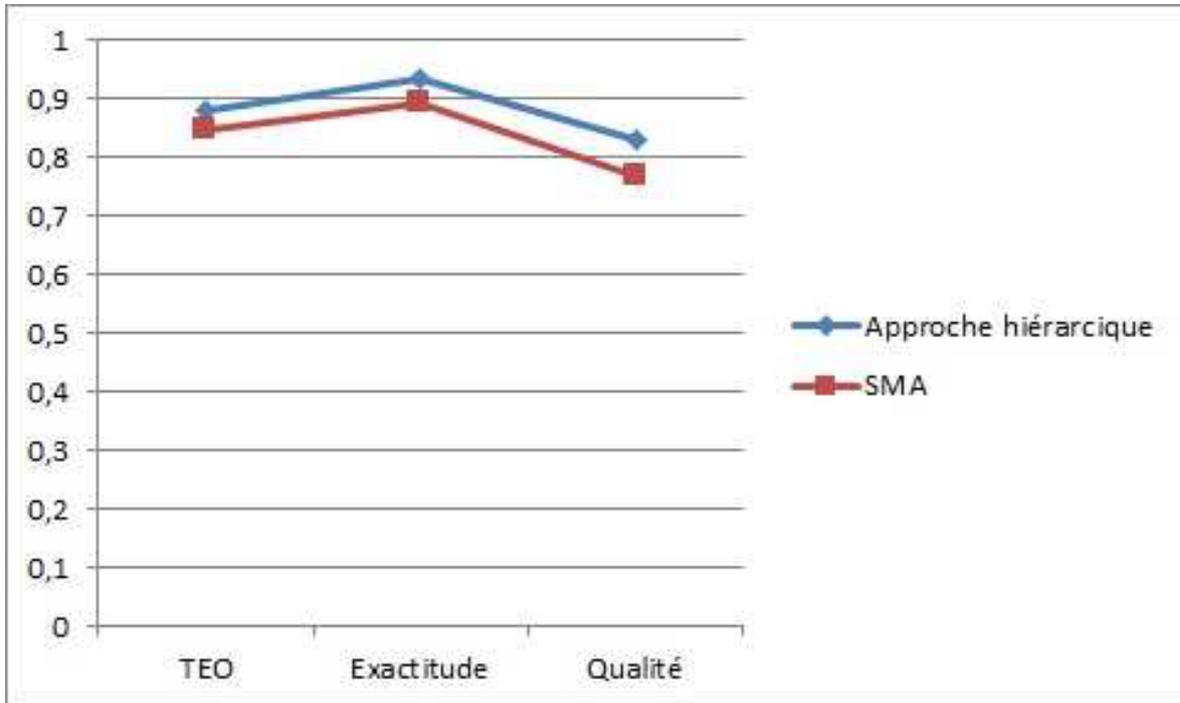


Figure 6.12 — Comparaison entre les performances des deux approches.

Tableau 6.4 — Performance des approches d'extraction de bâtiments

Approche	TEO	Exactitude
[Erener, 2013]	0.77	0.96
[Bouziani <i>et al.</i> , 2010]	0.92	0.9
[Stankov et He, 2013]	0.85	0.85
[Hu <i>et al.</i> , 2011]	0.86	0.88
[Lhomme <i>et al.</i> , 2009]	0.79	0.75
[Lari et Ebadi, 2007]	0.84	0.8
Template Dynamique	0.88	0.93
SMA	0.84	0.89

à très haute résolution. Nous remarquons que l'approche par template dynamique se positionne en deuxième position pour les deux facteurs ce qui indique que les résultats obtenus sont assez concurrentiels. Pour le système multi-agents que nous avons introduit, il se positionne parmi les quatre premières approches pour le facteur d'exactitude et avoisine la moyenne pour la TEO. Cependant, certaines des approches présentes dans cette comparaison utilisent des données exogènes comme [Bouziani *et al.*, 2010]

qui utilisent des cartes et des bases de données géographiques tandis que [Erener, 2013] utilise une image " pansharp " qui améliore la résolution.

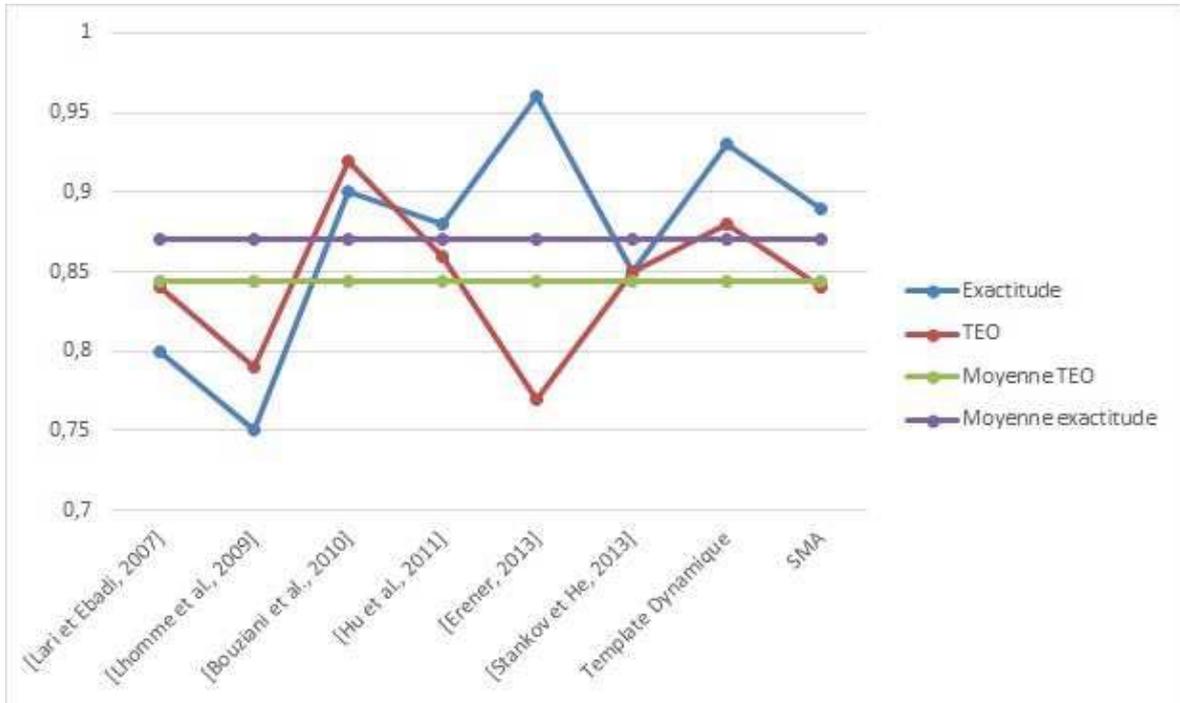


Figure 6.13 — Comparaison entre les performances de différentes approches d'extraction de bâtiments.

6.5 Évaluation de la classe Route

Nous présentons dans cette section les résultats d'extraction de la classe "Route".

Tableau 6.5 — Evaluation quantitative de la classe route

Sites	OCE	OPE	ONE	FA	TEO	Exactitude	Qualité %
site Cronenbourg ouest	92	8	6	5	86.67	94.84	82.88
site La Meinau	90	3	5	3	91.83	96.77	89.10
site Eckbolsheim	30	3	0	5	90.90	85.71	78.94
site Quartier des Quinze	20	8	0	7	71.42	74.07	57.14
site Oberhausbergen	26	2	0	7	92.85	78.78	74.28

Le tableau 6.5 expose trois métriques qui nous permettent de nous renseigner sur la performance de notre approche. La moyenne de la complétude (TEO), qui représente



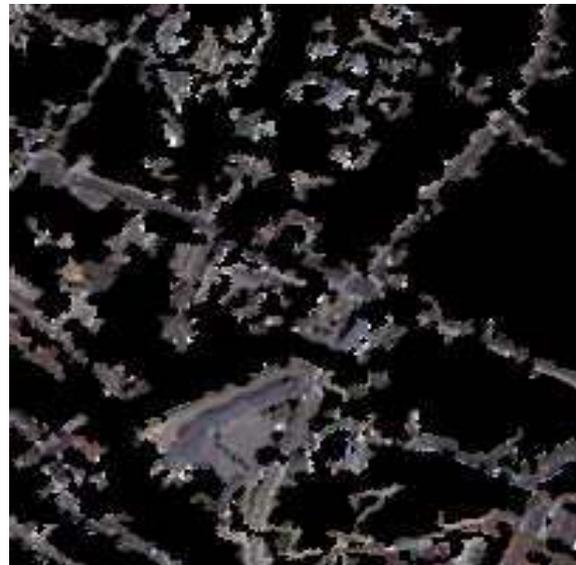
(a) Résultat d'extraction de l'approche hiérarchique du site de Cronenbourg ouest



(b) Résultat d'extraction de l'approche hiérarchique du site de La Meinau.

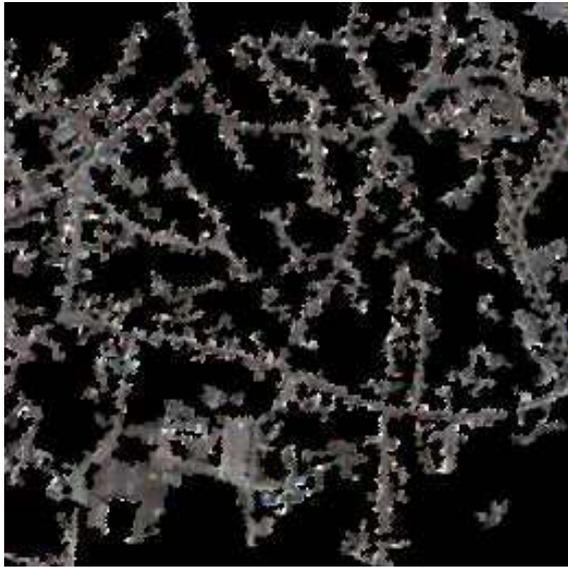


(c) Résultat d'extraction de l'approche hiérarchique du site d'Oberhausbergen.



(d) Résultat d'extraction de l'approche hiérarchique du site du Quartier des Quinze.

Figure 6.14 — Résultat de l'extraction des routes dans les images de test (...)



(a) Résultat d'extraction de l'approche hiérarchique du site d'Eckbolsheim.

Figure 6.15 — Résultat de l'extraction des routes dans les images de test (suite).

le facteur de rappel (i.e. nombre de segments routes correctement identifiés divisés par le nombre total de segments routes dans le réseau), est de l'ordre de 86 % ce qui implique une bonne localisation des routes permettant d'éviter la divergence dans les zones confuses (i.e. des objets ayant une radiométrie proche de celle des routes). Cependant, ce résultat illustre une étape de localisation et nécessite une phase de suivi permettant le franchissement d'obstacles isolés (c.f. voitures stationnées, arbres sur les bords, ombres). L'absence de l'étape de suivi, ainsi que la grande complexité du réseau routier et la densité des objets dans le site du Bas Rhin, explique entre autre le nombre élevé de portions de route partiellement extraites et non extraites ce qui a influé sur les résultats pour ce site.

En outre, la moyenne du facteur d'exactitude est de l'ordre de 86 %. Elle correspond au pourcentage des objets extraits et qui appartiennent à un voisinage de route. C'est à dire que peu d'éléments superflus ne correspondant pas au critère requis, sont détectés et retournés comme réponse du système. Ceci se répercute sur la précision de notre approche.

Concernant la moyenne du facteur de Qualité, elle est de l'ordre de 76 % impliquant une qualité assez acceptable d'extraction.



(a) Rond-point dans un site très dense

(b) Vérité terrain

Figure 6.16 — Image représentant un rond-point dans un site très dense.

Afin de valider la robustesse de notre approche par rapport aux formes circulaires des routes, nous avons testé notre approche sur les figures 6.16(a) et 6.17(a) qui représentent respectivement une zone très dense et contenant un petit rond-point et un tronçon de route circulaire (figure 6.16(a)) et la deuxième représentant un rond-point dans une zone ne contenant que de la végétation (figure 6.17(a)). Les figures 6.16(b) et 6.17(b) représentent la réalité terrain correspondante à ces deux figures.

La figure 6.18(a) illustre le résultat d'extraction de notre approche pour la zone dense (6.17(a)). Le rond-point est bien extrait ainsi que les routes composant l'image. Cependant, il reste des portions de route non détectées mais le résultat semble assez satisfaisant. La figure 6.18(b) représente le résultat d'extraction de notre approche appliquée sur la figure 6.17(a). Nous remarquons que le rond-point est bien extrait et que le résultat est encourageant.

Nous présentons dans le tableau 6.6 les résultats obtenus par quelques approches d'extractions de réseau routier afin de situer nos résultats (HCBRRRE) par rapport à ceux-ci. Nous remarquons que les résultats de notre approche restent satisfaisants malgré le fait que la majorité des approches auxquelles nous nous référons complètent l'étape d'extraction par un filtrage pour éliminer les faux segments et par une étape de suivi des routes pour regrouper les portions de routes. En particulier, notre approche se situe en deuxième position pour la TEO. Pour l'exactitude, notre approche



(a) Rond-point dans un site très dense

(b) Vérité terrain

Figure 6.17 — Image représentant un rond-point dans un site épars.



(a) Site dense

(b) Site épars

Figure 6.18 — Résultats d'extraction des ronds-point.

se positionne parmi les trois meilleures approches comparées (figure 6.19).

Tableau 6.6 — Performance des approches d'extraction de routes

Approche	TEO	Exactitude
[Clode <i>et al.</i> , 2004]	0.86	0.69
[Zhang et Couloigner, 2006b]	0.5	0.51
[Hu <i>et al.</i> , 2007]	0.85	0.91
[Naouai <i>et al.</i> , 2010]	0.82	0.76
[Zhang <i>et al.</i> , 2010]	0.79	0.81
[Naouai <i>et al.</i> , 2011]	0.85	0.93
HCBRRE	0.86	0.86

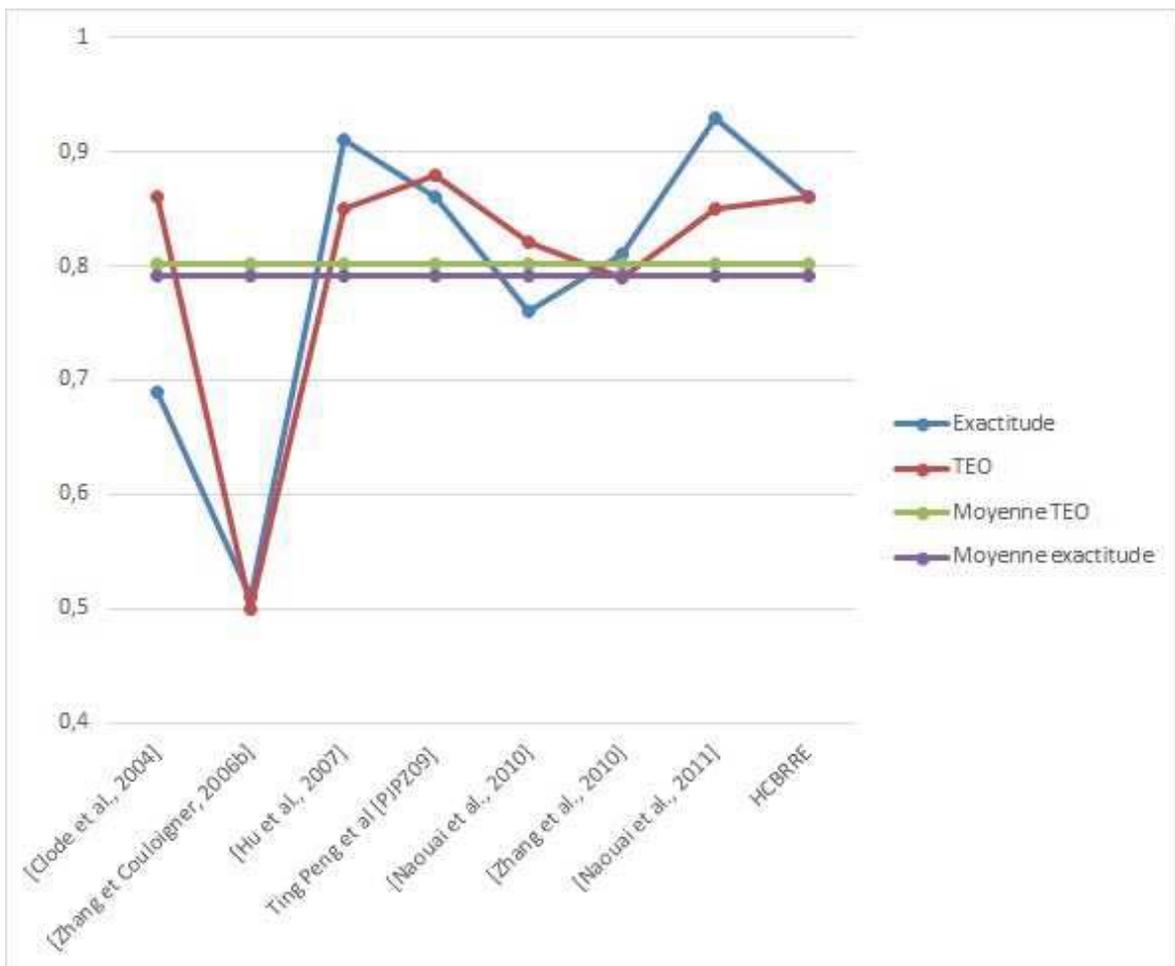


Figure 6.19 — Comparaison entre les performances de différentes approches d'extraction de routes.

6.6 Conclusion

Nous avons présenté dans ce chapitre une étude expérimentale des approches que nous avons introduit dans cette thèse. Les résultats obtenus montrent que les approches proposées sont assez satisfaisantes. Afin de valider la robustesse de ces approches, des images ayant différentes caractéristiques ont été testées. Nous avons aussi comparé nos travaux avec quelques travaux existants. Cette comparaison a montré que nos résultats sont concurrentiels et apparaissent dans les approches qui présentent les meilleurs taux d'extraction.

Conclusion générale et perspectives

Conclusion

Nous avons introduit dans cette thèse deux nouvelles approches d'analyse d'image orientées objets. Ces deux approches permettent de remédier à l'une des lacunes les plus importantes des approches objets, à savoir la dépendance de la phase de classification aux résultats de la phase de segmentation. Le manque d'interaction entre ces deux étapes est dû à la structure séquentielle des approches objets. L'idée était donc de proposer des approches qui permettent la collaboration entre les deux étapes de segmentation et de classification. Les deux approches proposées sont basées sur les deux techniques de segmentation les plus propices à la collaboration, à savoir, les techniques de segmentation par croissance de régions et les techniques de segmentation par collaboration région-contour.

La première approche proposée est une approche par croissance de région hiérarchique basée sur la classification. En combinant la croissance de région avec la classification, nous avons ciblé un des problèmes des approches OBIA, à savoir, le manque d'interaction entre la segmentation et la classification. Nous avons aussi visé les problèmes des algorithmes de croissance de régions, à savoir, le problème de la gestion des germes de départ, ainsi que le problème de choix des critères de croissance. En effet, Cette approche a permis de gérer le choix des germes de départ en se basant sur le score de similarité. Plus le score de similarité de la région augmente, plus la confiance en cette région augmente. Les germes pour chaque itération seront alors les régions non encore traitées en maximisant ce score de confiance. Concernant le choix des critères de croissance, il ne sera plus générique pour tous les germes de départ

mais spécifique à la classe du germe. Ceci permettra d'intégrer la sémantique dans le processus de croissance.

L'interaction entre le processus de segmentation et de classification est donc assurée grâce à une première classification qui guidera par la suite la segmentation par croissance de région. La mise en cause des régions non encore traitées est possible à travers la croissance des germes de chaque itération. Les régions déjà traitées formeront quant à elles des contraintes de croissance pour les germes non encore traités. Ceci permettra de diminuer l'espace de croissance.

La deuxième approche proposée est une approche multi agents évolutionnaire coopérative région-contour. Cette approche se base sur les systèmes multi-agents. Des agents région et contour sont modélisés afin de détecter les objets d'intérêts dans l'image. Le problème rencontré avec cette collaboration est le paramétrage des agents région et contour. Afin d'y remédier, nous avons proposé d'introduire une évolution locale pour chaque agent en se basant sur les algorithmes génétiques. Ceci a permis de gérer chaque objet d'une façon locale offrant ainsi une meilleure adaptation aux différents paramètres pouvant changer les propriétés intrinsèques de l'objet comme le bruit ou les occlusions.

Contributions et Résultats

Quelques contributions ont couronné ce travail de recherche. Une première contribution concerne l'introduction d'une approche hiérarchique. Cette approche, bien que dédiée à l'analyse des images satellitaires à très haute résolution, reste applicable dans plusieurs domaines de l'imagerie numérique. En effet l'approche hiérarchique offre une nouvelle vision pour la croissance de régions permettant une meilleure gestion de l'extraction des germes de départ et une intégration de la sémantique dans l'étape de croissance.

Aussi, nous avons proposé deux approches spécifiques aux classes route et bâtiment afin de valider l'apport de l'intégration de la sémantique dans le processus de croissance. L'apport concernant la classe bâtiment est basé sur l'intégration de templates dynamiques permettant de limiter l'espace de croissance. Une fois cet espace généré, une croissance fondée sur une fonction de fitness est appliquée afin d'extraire le bâtiment.

Concernant la classe route, nous nous sommes basés sur ses propriétés intrinsèques, à savoir que les routes sont rectilignes par morceau et qu'elles sont assez homogènes, cela est dû aux matériaux qui les composent. L'application de la transformée de Radon sur chaque germe permet d'extraire la direction de la route. Finalement la croissance est appliquée sur le critère d'homogénéité. Les résultats expérimentaux de la première approche sont très encourageants. Les tests ont été appliqués sur des images à très haute résolution spatiale de la zone de Strasbourg.

Ces différentes contributions de la première approche ont été validées par une revue internationale [Sellaouti *et al.*, 2014] et par des conférences internationales [Sellaouti *et al.*, 2012a, Sellaouti *et al.*, 2012c] et francophones [Sellaouti *et al.*, 2013a].

La deuxième partie concernant le système multi-agents évolutionnaire coopératif a donné naissance à deux contributions. La première, concerne le système multi-agents. La seconde, a introduit le paramétrage des agents qui est fait automatiquement en se basant sur un algorithme évolutionnaire local. Cette automatisation permet de gérer chaque objet séparément, d'offrir une grande flexibilité au système et d'assurer la gestion du bruit et des changements interclasses. L'évaluation de cette approche a été effectuée sur la classe "bâtiment" et l'algorithme génétique a été guidé par une fonction de fitness modélisant les propriétés de cette classe. Même si les résultats obtenus sont assez encourageants, l'approche reste basée sur une méta-heuristique qui ne cherche pas le meilleur résultat mais plutôt un résultat satisfaisant dans un temps raisonnable. Les résultats de la première approche sont meilleurs que ceux de l'approche évolutionnaire. Cette approche a été validée par une publication dans une revue [Sellaouti *et al.*, 2012b] et par une publication dans une conférence [Sellaouti *et al.*, 2013b] internationale.

Perspective

Même si ce travail présente des résultats assez encourageants en proposant des solutions pour la problématique abordée, plusieurs perspectives et améliorations sont envisagées. La première approche, proposant une vision itérative permettant une collaboration entre la classification et la segmentation, reste tributaire de la phase de prétraitement. En effet, si un germe est mal classé, alors la croissance se fera sur une base erronée et le résultat sera automatiquement biaisé. Ceci peut aussi influencer la croissance d'autres germes dans les itérations suivantes en imposant de fausses

contraintes. Une mise en cause de cette première phase de prétraitement semble essentielle. Deux pistes sont à étudier.

La première idée est de retravailler la première étape en se basant sur le résultat final de façon à réévaluer la classification obtenue. L'utilisation du méta-apprentissage peut permettre une ré-estimation de la segmentation et de la classification initiale. En effet, le choix des paramètres ou bien le choix des algorithmes utilisés dans cette étape peut influencer directement sur le résultat final.

La deuxième perspective se base sur le traitement des régions jugées mal classées. Ces régions peuvent être retraitées en les segmentant de nouveau et en évaluant une nouvelle fois le processus de croissance en prenant en compte les voisins de ces régions. Ceci peut permettre une réaffectation des segments composant la région traitée.

Bibliographie

- [Aksoy et Ercanoglu, 2012] AKSOY, B. et ERCANOGLU, M. (2012). Landslide identification and classification by object-based image analysis and fuzzy logic : An example from the azdavay region (kastamonu, turkey). *Computers & Geosciences*, 38(1):87 – 98.
- [Aminipouri et al., 2009] AMINIPOURI, M., SLIUZAS, R. et KUFFER, M. (2009). Object oriented analysis of very high resolution orthophotos for estimating the population of slum areas, case of dar es salaam, tanzania. *In Proceedings of the ISPRS conference :High-Resolution Earth Imaging for Geospatial Information :ISPRS XXXVIII, Hannover, Germany(2009)*.
- [Athanasiadis et Kollias, 2007] ATHANASIADIS, T. et KOLLIAS, S. (2007). A graph based, semantic region growing approach in image segmentation. *In Proceedings of 1st K-Space PhD Students Workshop, Berlin, Germany, September 2007*.
- [Athanasiadis et al., 2007] ATHANASIADIS, T., MYLONAS, P., AVRITHIS, Y. et KOLLIAS, S. (2007). Semantic image segmentation and object labeling. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(3):298 – 312.
- [Attarzadeh et Momeni, 2012] ATTARZADEH, R. et MOMENI, M. (2012). Object-based building extraction from high resolution satellite imagery. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B4:57–60.
- [Ayadi et al., 2010] AYADI, I., BOULLAUT, L., AKNIN, P. et SIARRY, P. (2010). Optimisation par algorithme génétique de la maintenance préventive dans un contexte de modélisation par modèles graphiques probabilistes. *In la Maitrise des RISQUES*, I. P., éditeur : *Actes du 17 ème Congrès de Maitrise des Risques et de Sureté de Fonctionnement*, page 10p, La Rochelle, France.

- [Bangalore *et al.*, 2011] BANGALORE, N., HASSAN, W., MITRA, B., BIRCH, P., YOUNG, R. et CHATWIN, C. (2011). Door surveillance using edge map-based harris corner detector and active contour orientation. *Proceedings of SPIE*, 8056(1):805608–805608–10.
- [Beasley *et al.*, 1993] BEASLEY, D., BULL, D. R. et MARTIN, R. R. (1993). An overview of genetic algorithms : Part 1, fundamentals.
- [Bendjoudi, 2002] BENDJOUDI, H. (2002). Le coefficient de compacité de gravelius : analyse critique d’un indice de forme des bassins versants. *Hydrological Sciences-Journal-des Sciences Hydrologiques*, 47(6).
- [Bezdek *et al.*, 1999] BEZDEK, J., KELLER, J., KRISNAPURAM, R. et PAL, N. R. (1999). *Fuzzy models and algorithms for pattern recognition and image processing*. Kluwer Academic Publishers, Massachusetts, États-Unis.
- [Blaschke, 2010] BLASCHKE, T. (2010). Object based image analysis for remote sensing. *ISPRS International Journal of Photogrammetry and Remote Sensing*, 65:2–16.
- [Blaschke et Lang, 2006] BLASCHKE, T. et LANG, S. (6-10 Novembre 2006). Object based image analysis for automated information extraction : A synthesis. *In Proceedings of MAPPs/ASPRS.*, San Antonia, Texas, Etats Unis.
- [Blaschke et Strobel, 2001] BLASCHKE, T. et STROBEL, J. (2001). What’s wrong with pixels? some recent developments interfacing remote sensing and gis. *GIS*, 6(1):12–17.
- [Bloch, 2002] BLOCH, I. (2002). Mathematical morphology and spatial relationships : Quantitative, semiquantitative and symbolic settings. *Applying Soft Computing in Defining Spatial Relationships*, pages 63–98.
- [Bouziani *et al.*, 2010] BOUZIANI, M., GOÏTA, K. et HE, D. (2010). Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 65(1):143 – 153.
- [Canny, 1986] CANNY, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679–714.
- [Chakroun *et al.*, 2011] CHAKROUN, M., WALI, A. et ALIM, A. M. (2011). Multi-agent system for moving object segmentation and tracking. *Advanced Video and Signal Based Surveillance, IEEE Conference on*, 0:424–429.

-
- [Chang *et al.*, 2011] CHANG, C., GORISSEN, B. et MELCHIOR, S. (2011). Fast oriented bounding box optimization on the rotation group. *ACM Trans. Graph.*, 30(5):122 :1 à 122 :16.
- [Clode *et al.*, 2004] CLODE, S., KOOTSOOKOS, P. et ROTTENSTEINER, F. (2004). The automatic extraction of roads from lidar data proceedings of american society of photogrammetry and remote sensing. *American Society of Photogrammetry and Remote Sensing*, 35:231–236.
- [Colliot *et al.*, 2005] COLLIOT, O., CAMARAY, O. et BLOCH, I. (2005). Un modèle déformable intégrant des relations spatiales pour la segmentation de structures cérébrales. *Information interaction intelligence*, 5(1):29–58.
- [Comer et Delp, 1999] COMER, M. L. et DELP, E. J. (1999). Morphological operations for color image processing.
- [Cover et Hart, 1967] COVER, T. et HART, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Information Theory*, 13(1):21–27.
- [Cui *et al.*, 2008] CUI, W., GUANAND, Z. et ZHANG, Z. (2008). An improved region growing algorithm for image segmentation. In *Proceedings of the 2008 International Conference on Computer Science and Software Engineering - Volume 06*, pages 93–96, Washington, DC, USA. IEEE Computer Society.
- [Dasarathy, 1990] DASARATHY, B. (1990). Nearest neighbor (nn) norms : Nn pattern classification techniques. IEEE Computer Society Press, Los Alamitos, Californie, États Unis.
- [Dehavri et Heck, 2009] DEHAVRI, A. et HECK, R. (2009). Comparison of object-based and pixel based infrared airborne image classification methods using dem thematic layer. *Journal of Geography and Regional Planning*, 2(4):86–96.
- [Dennis *et al.*, 2012] DENNIS, D. C., FRANKLIN, S. E. et DUBÉ, M. G. (2012). Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. *International Journal of Remote Sensing*, 33(14):4502–4526.
- [Deriche, 1987] DERICHE, R. (1987). Using canny’s criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1: 167–187.

- [Derivaux, 2009] DERIVAUX, S. (2009). *Construction et classification d'objets à partir d'images de télédétection par une approche itérative guidée par des connaissances du domaine*. Thèse de doctorat, Université de Strasbourg.
- [Derivaux et al., 2007] DERIVAUX, S., DURAND, N. et WEMMERT, C. (2007). On the complementarity of an ontology and a nearest neighbour classifier for remotely sensed image interpretation. *In Proceedings of IGARSS*, pages 3983–3986. IEEE.
- [Derivaux et al., 2010] DERIVAUX, S., FORESTIER, G., WEMMERT, C. et LEFBVRE, S. (2010). Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation. *Pattern Recognition Letters*, 31(15):2364–2374.
- [Deruyver et Hodé, 1997] DERUYVER, A. et HODÉ, Y. (1997). Constraint satisfaction problem with bilevel constraint : application to interpretation of over-segmented images. *Artificial Intelligence*, 93:321–335.
- [Deruyver et al., 2011] DERUYVER, A., HODÉ, Y. et JOLION, J. M. (2011). Graph consistency checking : a tool to check the semantic consistency of a segmentation. *Int. J. Semantic Computing*, 5(2):179–210.
- [Devi et Krishna, 2012] DEVI, Y. et KRISHNA, I. (2012). Pixel-based and object-oriented classification of high resolution satellite images. *Canadian Journal on Electrical and Electronics Engineering*, 3(1):31–34.
- [Dhiaf et Sellaouti, 2009] DHIAF, Z. B. et SELLAOUTI, A. (2009). Système de classification d'images forestières haute résolution basé sur l'intégration des informations et des connaissances structurelles. *In Proceedings of TAIMA, Hammamet, Tunisia*.
- [Diday et al., 1982] DIDAY, E., LEMAIRE, J. et POUGET, J. F. (1982). *Éléments d'analyse de données*. Dunod, Paris, France.
- [Do et al., 2011] DO, A. L., SENAME, O. et DUGARD, L. (2011). Optimisation par Algorithme Génétique d'une Commande LPV de Suspension Semi-active. *In Actes des Journées Doctorales Journées Nationales MACS (JD-JN- MACS 2011)*, page s/n, Marseille, France. Département Automatique Département Automatique.
- [Duda et Hart, 1973] DUDA, R. O. et HART, P. E. (1973). *Pattern Classification and Scene Analysis*. John Wiley and Sons, Chichester, Royaume Uni.
- [Durand et al., 2007] DURAND, N., DERIVAUX, S., FORESTIER, G., WEMMERT, C., GANÇARSKI, P., BOUSSAÏD, O. et PUISSANT, A. (2007). Ontology based object recognition for remote sensing image interpretation. *In IEEE International Confe-*

- rence on Tools with Artificial Intelligence*, volume 1, pages 472–479, Patras, Greece. IEEE Computer Society.
- [Erener, 2013] ERENER, A. (2013). Classification method, spectral diversity, band combination and accuracy assessment evaluation for urban feature detection. *International Journal of Applied Earth Observation and Geoinformation*, 21(0):397 – 408.
- [Ericksson, 2004] ERICKSSON, M. (2004). Species classification of individually segmented tree crowns in high resolution aerial images using radiometric and morphologic image measures. *Remote Sensing of Environment*, 91:469–477.
- [Eriksson, 2005] ERIKSSON, M. (2005). Comparison of three individual tree crown detection methods. *Machine Vision and Applications*, 16(4):258–265.
- [Fan *et al.*, 2001a] FAN, J., YAU, D. K. Y., ELMAGARMID, A. K. et AREF, W. (2001a). Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Trans. On Image Processing*, 10:1454–1466.
- [Fan *et al.*, 2005] FAN, J., ZENG, G., BODY, M. et HACID, M. S. (2005). Seeded region growing : an extensive and comparative study. *Pattern Recognition Letters*, pages 1139–1156.
- [Fan *et al.*, 2001b] FAN, J., ZHU, X. et WU, L. (2001b). Automatic model-based semantic object extraction algorithm. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11(10):1073 –1084.
- [Fekir et Benamrane, 2011] FEKIR, A. et BENAMRANE, N. (2011). Segmentation of medical image sequence by parallel active contour. *Advances in experimental medicine and biology*, 696:515–522. PMID : 21431592.
- [Ferber, 1995] FERBER, J. (1995). *Les systèmes multi-agents : vers une intelligence collective*. Informatique, Intelligence Artificielle. Inter Editions.
- [Ferber et Gasser, 1991] FERBER, J. et GASSER, L. (1991). *Expert systems & their applications*. Numéro 9. conferences/ tutorials/exhibition, Eleventh International Workshop, Avignon, May 27-31, 1991. Intelligence artificielle distribuée / Jacques Ferber & Les Gasser.
- [Fischler *et al.*, 1981] FISCHLER, M., TENENBAUM, J. et WOLF, H. C. (1981). Detection of roads and linear structures in low-resolution aerial imagery using a multi-source knowledge integration technique. *Computer Graphics and Image Processing*, 15:201–223.

- [Fleureau *et al.*, 2009] FLEUREAU, J., GARREAU, M., BOULMIER, D., LECLERCQ, C. et HERNANDEZ, A. I. (2009). Segmentation 3D multi-objets d'images scanner cardiaques : une approche multi-agents. *IRBM*, 30(3):104–113. Ce travail a été financé par le projet européen ALFA IPECA et supporté par le ministère de la Recherche en France (projet CITH no 04T187-188-189-190).
- [Forestier, 2010] FORESTIER, G. (Septembre, 2010). *Connaissances et clustering collaboratif d'objets complexes multisources*. Thèse de doctorat, Université de Strasbourg, France.
- [Forestier *et al.*, 2008a] FORESTIER, G., DERIVAUX, S., WEMMERT, C. et GANÇARSKI, P. (Boston, Massachusetts, july 2008a). An evolutionary approach for ontology driven image interpretation. In *Proceedings of EvoWorkshops*, volume 4974 de *Lecture Notes in Computer Science*, pages 295–304. Springer.
- [Forestier *et al.*, 2012] FORESTIER, G., PUISSANT, A., WEMMERT, C. et GANÇARSKI, P. (2012). Knowledge-based region labeling for remote sensing image interpretation. *Computers, Environment and Urban Systems*, 36(5):470 – 480.
- [Forestier *et al.*, 2008b] FORESTIER, G., WEMMERT, C. et GANÇARSKI, P. (2008b). Multisource images analysis using collaborative clustering. *EURASIP J. Adv. Signal Process*, 2008:133 :1–133 :11.
- [Forestier *et al.*, 2008c] FORESTIER, G., WEMMERT, C. et GANÇARSKI, P. (2008c). On combining unsupervised classification and ontology knowledge. In *Proceedings of IEEE International Geoscience & Remote Sensing Symposium, IGARSS, Boston, Massachusetts, USA*, pages 395–398.
- [Forgy, 1965] FORGY, E. (1965). Cluster analysis of multivariate data : efficiency versus interpretability of classification. *Biometrics num 21*, 21:768–780.
- [Gao et Wu, 2004] GAO, J. et WU, L. (12-23 July 2004). Automatic extraction of road networks in urban areas from ikonos imagery based on spatial reasoning. In *Proceedings of XXth ISPRS Congress, The international Archives of the Photogrammetry Remote Sensing and Spatial information Sciences*.
- [Gao *et al.*, 2011] GAO, Y., MARPU, P., NIEMEYER, I., RUNFOLA, D., GINER, N. et PONTIUS, T. H. R. G. (2011). Object based classification with features extracted by a semi automatic feature extraction algorithm seath. *Geocarto International*, 26(3):211–226.

- [Gao et Mas, 2008] GAO, Y. et MAS, J. F. (2008). A comparison of the performance of pixel based and object based classifications over images with various spatial resolutions. *Online Journal of Earth Sciences*, 2.
- [Ghandour, 2010] GHANDOUR, S. (2010). *Segmentation d'images couleurs par morphologie mathématique : application aux images microscopiques*. Thèse de doctorat, Génie électrique, électronique, télécommunications.
- [Gholoobi et al., 2010] GHOLOOBI, M., TAYYEBI, A., TALEYI, M. et TAYYEBI, A. H. (2010). Comparing pixel based and object based approaches in land use classification in mountainous areas. *In Proceedings of XXXVIIIth ISPRS Congress, International archives of the Photogrammetry, Remote Sensing and Spatial Information Science*.
- [Giraudon et al., 1992] GIRAUDON, G., GARNESSON, P. et MONTÉSINON, P. (1992). Messie : a multi expert system in computer vision. application for aerial imagery interpretation. *Traitement du signal*, 9(5):403–419.
- [Girgis et al., 2009] GIRGIS, M. R., SEWISY, A. A. et MANSOUR, R. F. (2009). A robust method for partial deformed fingerprints verification using genetic algorithm. *Expert Syst. Appl.*, 36(2):2008–2016.
- [Grigillo et Kanjir, 2012] GRIGILLO, D. et KANJIR, U. (2012). Urban object extraction from digital surface model and digital aerial images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-3:215–220.
- [Haralick, 1979] HARALICK, R. (Mai 1979). Statistical and structural approaches to texture. *PROCEEDINGS OF THE IEEE*, 67(5).
- [Harris et Stephens, 1988] HARRIS, C. et STEPHENS, M. (1988). A combined corner and edge detection. *In Proceedings of The Fourth Alvey Vision Conference*, pages 147–151.
- [Hay et Castilla, 2006] HAY, G. J. et CASTILLA, G. (July 4-5, 2006). Object-based image analysis : Strengths, weaknesses, opportunities and threats (swot). *In Proceedings of 1st International Conference on Object-based Image Analysis (OBIA 2006)*.
- [Hentati et al., 2011] HENTATI, J., NAOUAI, M., HAMOUDA, A. et WEBER, C. (2011). Measuring rectangularity using gr-signature. *In Proceedings of the Third Mexican conference on Pattern recognition, MCPR'11*, pages 136–145, Berlin, Heidelberg. Springer-Verlag.

- [Hofmann, 2001] HOFMANN, P. (2001). Detecting urban features from ikonos data using an object-oriented approach. *Remote Sensing Photogrammetry Society*, 1:28–33.
- [Holland, 1962] HOLLAND, J. H. (1962). Outline for a logical theory of adaptive systems. *J. ACM*, 9(3):297–314.
- [Hough, 1962] HOUGH, P. (1962). Methods and means for recognizing complex patterns. *U.S. Patent*.
- [Hu et al., 2007] HU, J., RAZDAN, A., FEMIANI, J. C., CUI, M. et WONKA, P. (2007). Road network extraction and intersection detection from aerial images by tracking road footprints. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(12):4144–4157.
- [Hu et al., 2011] HU, L., ZHENG, J. et GAO, F. (2011). A building extraction method using shadow in high resolution multispectral images. *In Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 1862–1865.
- [Inglada et Michel, 2009] INGLADA, J. et MICHEL, J. (2009). Qualitative spatial reasoning for high-resolution remote sensing image analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(2):599–612.
- [Jacquin et al., 2008] JACQUIN, A., MISAKOVA, L. et GAY, M. (2008). A hybrid object-based classification approach for mapping urban sprawl in periurban environment. *Landscape and Urban Planning*, 84(2):152–165.
- [Jain et al., 1999] JAIN, A., MURTY, M. et FLYNN, P. (1999). Data clustering : A review. *ACM Computing Surveys*, 31:264–323.
- [Jaynes et al., 1994] JAYNES, C., STOLLE, F. et COLLINS, R. (1994). Task driven perceptual organization for extraction of rooftop polygons. *In Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, pages 152–159.
- [Jennings et al., 2001] JENNINGS, N., FARATIN, P., LOMUSCIO, A., PARSONS, S., WOOLDRIDGEAND, M. et C.SIERRA (2001). Automated negotiation : prospects, methods and challenges. *Group Decision and Negotiation*, 10(2):199–215.
- [Jiang et al., 2011] JIANG, B., QIAN, M., NIU, L., SONG, R., ZHU, H., JIN, Q. et ZHENG, H. (2011). Region growing based on window frame difference for segmenting ultrasonic contrast images of arteries. *In Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference on*, pages 1–4.

- [Kang *et al.*, 2011] KANG, J., XIAO, C., DENG, M. et LIU, J. Y. H. (2011). Image registration based on harris corner and mutual information. *In Electronic and Mechanical Engineering and Information Technology (EMEIT), 2011 International Conference on*, volume 7, pages 3434–3437.
- [Karsenty *et al.*, 2007] KARSENTY, A., ANTUNES, A. et CENTENO, J. (Avril, 2007). Classification orientée objet de la perméabilité des sols en zone urbaine à l’aide d’imagerie très haute résolution et de données laser scanner à curitiba. *In Proceedings of Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Brésil*, pages 562–572.
- [Kass *et al.*, 1988] KASS, M., WITKIN, A. et TERZOPOULOS, D. (1988). Snakes : Active contour models. *International Journal of Computer Vision*, pages 321–331.
- [Kurtz, 2012] KURTZ, C. (2012). *Approche collaborative segmentation classification pour l’analyse descendante d’images multirésolutions*. Thèse de doctorat, Université de Strasbourg.
- [Lang *et al.*, 2009] LANG, S., ALBRECHT, F. et BLASCHKE, T. (2009). *OBIA Tutorial : Introduction to Object-based Image Analysis*. Salzburg, v 1.0 édition.
- [Lari et Ebadi, 2007] LARI, Z. et EBADI, H. (2007). Automated building extraction from high-resolution satellite imagery using spectral and structural information based on artificial neural networks. *In Proceedings of the ISPRS Hannover Workshop 2007 High-Resolution Earth Imaging for Geospatial Information, Hannover, Germany(2007)*.
- [Lefebvre *et al.*, 2011] LEFEBVRE, A., CORPETTI, T. et HUBERT-MOY, L. (2011). On-delettes et théorie des évidences pour la classification orientée-objet. caractérisation et suivi des changements d’occupation des sols de la métropole de rennes. *Revue Internationale de Géomatique*, 21(3):297–325.
- [Lhomme *et al.*, 2009] LHOMME, S., HE, D., WEBER, C. et MORIN, D. (2009). A new approach to building identification from very-high-spatial-resolution images. *International Journal of Remote Sensing*, 30.
- [Mahmoudi *et al.*, 2013] MAHMOUDI, F. T., SAMADZADEGAN, F. et REINARTZ, P. (2013). Object oriented image analysis based on multi-agent recognition system. *Computers & Geosciences*.
- [Maillot, 2005] MAILLOT, N. (2005). *Ontology Based Object Learning and Recognition*. Phd thesis, Université de Nice Sophia Antipolis.

- [Maillot *et al.*, 2003] MAILLOT, N., THONNAT, M. et BOUCHER., A. (2003). Towards ontology based cognitive vision. *In Proceedings of ICVS*, pages 44–53.
- [Marangoz *et al.*, 2004] MARANGOZ, A., ORUC, M. et BUYUKSALIH, G. (July, 2004). Object-oriented image analysis and semantic network for extracting the roads and buildings from ikonos pan-sharpened images. *In Proceedings of the ISRPS 2004 Annual Conference, Istanbul, Turquie*, pages 19–23.
- [Mather, 1987] MATHER, P. M. (1987). *Computer Processing of Remotely-Sensed Images. An Introduction*. John Wiley and Sons, Chichester, Royaume Uni.
- [Matinfar *et al.*, 2007] MATINFAR, H., SARMADIAN, F., PANAH, S. A. et HECK, R. (2007). Comparisons of object-oriented and pixel-based classification of land use/land cover types based on landsat7, etm+ spectral bands (case study : Arid region of iran). *american eurasian journal agricultural environmental science*, 4(2):448–456.
- [Mayunga *et al.*, 2007] MAYUNGA, S. D., COLEMAN, D. J. et ZHANG, Y. (2007). A semi-automated approach for extracting buildings from quickbird imagery applied to informal settlement mapping. *Int. J. Remote Sens.*, 28(10):2343–2357.
- [Mazouzi, 2008] MAZOUZI, S. (2008). *Reconnaissance de Formes par les Systèmes Auto-Organisme*. Thèse de doctorat, Université Mentouri De Constantine Faculté Des Science De L’ingénieur Département D’informatique.
- [Medina-Carnicer *et al.*, 2011] MEDINA-CARNICER, R., MUNOZ-SALINAS, R., YEGUAS-BOLIVAR, E. et L.DIAZ-MAS (2011). A novel method to look for the hysteresis thresholds for the canny edge detector. *Pattern Recognition*, 44(6):1201–1211.
- [Mehnert et Jackway, 1997] MEHNERT, A. et JACKWAY, P. (1997). An improved seeded region growing algorithm. *Pattern Recogn. Lett.*, 18(10):1065–1071.
- [Melkemi *et al.*, 2006] MELKEMI, K. E., BATOUCHE, M. et FOUFOU, S. (2006). A multiagent system approach for image segmentation using genetic algorithms and extremal optimization heuristics. *Pattern Recognition Letters*, 27(11):1230–1238.
- [Mendes *et al.*, 2009] MENDES, J., GONSLVES, J. et RESENDE, M. (2009). A random key based genetic algorithm for the resource constrained project scheduling problem. *Computers amp ; Operations Research*, 36(1):92 – 109. Part Special Issue : Operations Research Approaches for Disaster Recovery Planning.

- [Merlet et Zerubia, 1996] MERLET, N. et ZERUBIA, J. (1996). New prospects in line detection by dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:426–431.
- [Middelkoop *et al.*, 1989] MIDDELKOOP, H., MILTENBURG, J. W. et MULDE, N. J. (1989). Knowledge engineering for image interpretation and classification : a trial run. *ITC Journal*, 1:27–32.
- [Moussa, 2011] MOUSSA, R. (2011). *Ségmentation Multi-agents en Imagerie Biologique Et Médicale : Application Aux IRM 3D*. These, Université Bordeaux I; Université Sciences et Technologies - Bordeaux I.
- [Myint *et al.*, 2011] MYINT, S. W., GOBER, P., BRAZEL, A., GROSSMAN-CLARKE, S. et WENG, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115(5):1145 – 1161.
- [Nacereddine *et al.*, 2010] NACEREDDINE, N., TABBONE, S., ZIOU, D. et HAMAMI, L. (2010). Shape-based image retrieval using a new descriptor based on the radon and wavelet transforms. *In In proceedings of Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1997–2000.
- [Naouai *et al.*, 2011] NAOUAI, M., HAMOUDA, A., NARJESS, M. et WEBER, C. (2011). Linear structure recognition based on image vectorization. *In CSURKA, G., KRAUS, M. et BRAZ, J., éditeurs : IMAGAPP/IVAPP*, pages 53–59. SciTePress.
- [Naouai *et al.*, 2010] NAOUAI, M., HAMOUDA, A. et WEBER, C. (2010). Urban road extraction from high-resolution optical satellite images. *In Proceedings of International Conference on Image Analysis and Recognition*, pages 420–433.
- [Nelson *et al.*, 2009] NELSON, A. L., BARLOW, G. J. et DOITSIDIS, L. (2009). Fitness functions in evolutionary robotics : A survey and analysis. *Robotics and Autonomous Systems*, 57(4):345 – 370.
- [Ortner, 2004] ORTNER, M. (2004). *Processus ponctuels marqués pour l'extraction automatique de caricatures de bâtiments à partir de modèles numériques d'élévation*. These, Université Nice Sophia Antipolis.
- [Pal et Pal, 1993] PAL, N. R. et PAL, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition*, pages 1277–1294.
- [Pei *et al.*, 2010] PEI, Y., WU, H., YU, J. et CAI, G. (2010). Effective image registration based on improved harris corner detection. *In Information Networking and*

- Automation (ICINA), 2010 International Conference on*, volume 1, pages V1–93–V1–96.
- [Plataniotis et Venetsanopoulos, 2000] PLATANIOTIS, K. N. et VENETSANOPOULOS, A. N. (2000). *Color image processing and applications*.
- [Poulain, 2010] POULAIN, V. (2010). *Fusion d'images optique et radar à haute résolution pour la mise à jour de bases de données cartographiques*. Phd thesis, National Polytechnic Institute of Toulouse, Toulouse, France.
- [Puissant, 2003] PUISSANT, A. (2003). *Information géographique et image à Très Haute Résolution*. Thèse de doctorat, Université Strasbourg 1, France.
- [QIAN *et al.*, 2007] QIAN, J., ZHOU, Q. et HOU, Q. (28-29 August 2007). Comparison of pixel-based and object-oriented classification methods for extracting built-up areas in aridzone. *In Proceedings of ISPRS Workshop on Updating Geo-spatial Databases with Imagery & The 5th ISPRS Workshop on Dynamic and Multi-dimensional GIS*, Urumchi, China.
- [Radon, 1917] RADON, J. (1917). Ufiber die bestimmung von funktionen durch ihre integral-werte langs gewisser mannigfaltigkeiten. *Berichte Sachsische Akademie der Wissenschaften, Leipzig, Math-Phys. Kl.*, 62:262–267.
- [Reyes-Aldasoro et Bhalerao, 2006] REYES-ALDASORO, C. et BHALERAO, A. (2006). The bhattacharyya space for feature selection and its application to texture segmentation. *Pattern Recogn.*, 39(5):812–826.
- [Robert *et al.*, 2010] ROBERT, C., WEIH, J., NORMAN, D. et RIGGAN, J. (29 June - 2 July, 2010). Object-based classification vs. pixel-based classification : Comparative importance of multi-resolution imagery. *In Proceedings of XXXVIIIth ISPRS Congress, The international Archives of the Photogrammetry Remote Sensing and Spatial information Sciences*.
- [Rochery *et al.*, 2005] ROCHERY, M., JERMYN, I. et ZERUBIA, J. (2005). Higher order active contours. *International Journal of Computer Vision*.
- [Roelfsema *et al.*, 1998] ROELFSEMA, P. R., LAMME, V. A. F. et SPEKREIJSE, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395(6700):376–381.
- [Rosin, 1999] ROSIN, P. (1999). Measuring rectangularity. *Machine Vision and Applications*, 11:191–196.

- [Salehi *et al.*, 2011] SALEHI, B., ZHANG, Y. et ZHANG, M. (1-5 Mai 2011). Combination of object-based and pixel-based image analysis for classification of vhr imagery over urban areas. *In ASPRS 2011 Annual Conference*.
- [Sebari, 2008] SEBARI, I. (2008). *Développement d'une nouvelle approche basée objets pour l'extraction automatique de l'information géographique en milieu urbain à partir des images à très haute résolution spatiale*. Thèse de doctorat, Université de Sherbrooke, Québec, Canada.
- [Sebari et He, 2007] SEBARI, I. et HE, D. C. (2007). Les approches de segmentation d'image par coopération régions-contours. *Revue Télédétection*, 7(4):499–506.
- [Seema, 2012] SEEMA, J. (2012). Exploring the potential of object based image analysis for mapping urban land cover. *Journal of the Indian Society of Remote Sensing*, 40:507–518.
- [Sellaouti, 2009] SELLAOUTI, A. (2009). Intégration des informations et des connaissances structurelles dans la classification des images satellitaires forestières haute résolution. Mastère, Faculté des Sciences de Tunis, Tunisie.
- [Sellaouti *et al.*, 2012a] SELLAOUTI, A., HAMOUDA, A., DERUYVER, A. et WEMMERT, C. (2012a). Hierarchical classification-based region growing (hcbg) : A collaborative approach for object segmentation and classification. *In Image Analysis and Recognition - 9th International Conference, ICIAR 2012, Aveiro, Portugal.*, pages 51–60.
- [Sellaouti *et al.*, 2014] SELLAOUTI, A., HAMOUDA, A., DERUYVER, A. et WEMMERT, C. (2014). Template-based hierarchical building extraction. *Geoscience and Remote Sensing Letters, IEEE*, 11(3):706–710.
- [Sellaouti *et al.*, 2013a] SELLAOUTI, A., HAMOUDA, A., DERUYVER, A. et WEMMERT, C. (Janvier, 2013a). Approche orientée objet sémantique et coopérative pour la classification des images de zones urbaines à très haute résolution. *In Proceedings of 13e Conférence Francophone sur l'Extraction et la Gestion des Connaissances (EGC2013), Toulouse, France*, pages 103–114.
- [Sellaouti *et al.*, 2012b] SELLAOUTI, A., OULEDSGHAIER, M. et HAMOUDA, A. (2012b). An edge-region cooperative multi-agent approach for buildings extraction. *In Proceedings of the The Second International Conference on Digital Information and Communication Technology and its Applications(DICTAP2012)*, volume 7324, pages 51–60.

- [Sellaouti *et al.*, 2013b] SELLAOUTI, A., OULEDSGHAIER, M., HAMOUDA, A., DERUYVER, A. et WEMMERT, C. (2013b). Local evolutionary multi-agent system for buildings extraction. *Journal of the Indian Society of Remote Sensing*.
- [Sellaouti *et al.*, 2012c] SELLAOUTI, A., ROJBANI, H., HAMOUDA, A., DERUYVER, A. et WEMMERT, C. (Decembre, 2012c). Hierarchical classification-based radon road extraction (hcbrr). *In Proceedings of International Conference on Control, Automation, Robotics and Vision (ICARCV), Guanzhou, China*, pages 390–395.
- [Shih et Cheng, 2005] SHIH, F. et CHENG, S. (2005). Automatic seeded region growing for color image segmentation. *Image Vision Comput.*, 23:877–886.
- [Sibil *et al.*, 2010] SIBIL, A., GODIN, N., R’MILI, M. et FANTOZZI, G. (2010). Optimisation de la segmentation de données d’émission acoustique à l’aide d’un algorithme génétique. *In d’Acoustique SFA, S. F., éditeur : 10^{ème} Congrès Français d’Acoustique*, pages –, Lyon, France.
- [Silva *et al.*, 2010] SILVA, C., CENTENO, J. S. et HENRIQUES, M. (2010). Automatic road extraction in rural areas, based on the radon transform using digital images. *Canadian Journal of Remote Sensing*, 36(6):737–749.
- [Singh *et al.*, 2012] SINGH, D., MAURYA, R., SHUKLA, A., SHARMA, M. et GUPTA, P. (2012). Building extraction from very high resolution multispectral images using ndvi based segmentation and morphological operators. *In Proceedings of Engineering and Systems (SCES), 2012*, pages 1–5.
- [Sivaraj et Ravichandran, 2011] SIVARAJ, R. et RAVICHANDRAN, T. (2011). A review of selection methods in genetic algorithm. *International Journal of Engineering Science and Technology (IJEST)*, 3(5):3792–3797.
- [Sledge *et al.*, 2011] SLEDGE, I., KELLER, J., SONG, W. et DAVIS, C. (2011). Conflation of vector buildings with imagery. *Geoscience and Remote Sensing Letters, IEEE*, 8(1):83–87.
- [Smara *et al.*, 2003] SMARA, Y., OUARAB, N., LAAMA, S. et CHERIFI, D. (2003). Techniques de fusion et de classification floue d’images satellitaires multi sources pour la caractérisation et le suivi de l’extension du tissu urbain de la région d’alger. 2nd FIG Regional Conference, Marrakech, Maroc.
- [Smeulders *et al.*, 2000] SMEULDERS, A., WORRING, M., SANTINI, S., GUPTA, A. et JAIN, R. (2000). Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349–1380.

- [Stankov et He, 2013] STANKOV, K. et HE, D. C. (2013). Building detection in very high spatial resolution multispectral images using the hit-or-miss transform. *Geoscience and Remote Sensing Letters, IEEE*, 10(1):86 –90.
- [Suzuki, 2002] SUZUKI, H. (2002). *Exploitation de connaissances structurelles en classification d'images : une approche par partition flou et optimisation combinatoire*. Thèse de Doctorat, Université Paul Sabatier, Toulouse, France.
- [Tabbone et al., 2008] TABBONE, S., TERRADES, O. et BARRAT, S. (2008). Histogram of radon transform. a useful descriptor for shape retrieval. *In In proceeding of Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4.
- [Thibault et al., 2007] THIBAUT, G., DEVIC, C., FERTIL, B., MARI, J. et SEQUEIRA, J. (2007). Indices de formes : de la 2d vers la 3d application au classement de noyaux de cellules. Rapport technique.
- [Tinghua et al., 2013] TINGHUA, A., XIAOQIANG, C., PENGCHENG, L. et MIN, Y. (2013). A shape analysis and template matching of building features by the fourier transform method. *Computers, Environment and Urban Systems*, 41(0):219 – 233.
- [Toft, 1996] TOFT, P. (1996). *The Radon Transform, Theory and Implementation*. Thèse de doctorat, Department of Mathematical Modelling Section for Digital Signal Processing Technical University of Denmark.
- [Toussaint, 1983] TOUSSAINT, G. T. (1983). Solving geometric problems with the rotating calipers. *In Proceedings of IEEE MELECON*.
- [Vincent et Soille, 1991] VINCENT, L. et SOILLE, P. (1991). Watersheds in digital spaces : An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(6):583–598.
- [Walter, 2004] WALTER, V. (2004). Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4):225–238.
- [Wang et Li, 2006] WANG, J. et LI, Y. (2006). *MATLAB 7.0 Image Processing*. Beijing : National Defence Industry Publishing.
- [Y. Gao et al., 2011] Y. GAO, J. F. M., N.KERLE et PACHECO, J. A. N. (2011). Optimal region growing segmentation and its effect on classification accuracy. *Int. J. Remote Sens.*, 32:3747–3763.

- [Y. Zhang, 2004] Y. ZHANG, R. W. (12-23 July 2004). Multi-resolution and multispectral image fusion for urban object extraction. *In Proceedings of XXth ISPRS Congress, The international Archives of the Photogrammetry Remote Sensing and Spatial information Sciences*, Istanbul, Turkey.
- [Yang et Lee, 1997] YANG, H. et LEE, S. U. (1997). Split-and-merge segmentation employing thresholding technique. *In Proceedings of the 1997 International Conference on Image Processing (ICIP '97) 3-Volume Set-Volume 1 - Volume 1*, ICIP '97, pages 239–, Washington, DC, USA. IEEE Computer Society.
- [Yu et Clausi, 2008] YU, Q. et CLAUSI, D. A. (2008). Irgs : Image segmentation using edge penalties and region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(12): 2126–2139.
- [Zhang *et al.*, 2010] ZHANG, L., ZHANG, J., ZHANG, D., HOU, X. et YANG, G. (2010). Urban road extraction from high-resolution remote sensing images based on semantic model. *In Geoinformatics'10*, pages 1–5.
- [Zhang et Couloigner, 2006a] ZHANG, Q. et COULOIGNER, I. (2006a). Comparing different localization approaches of the radon transform for road centerline extraction from classified satellite imagery. *In ICPR (2)'06*, pages 138–141.
- [Zhang et Couloigner, 2006b] ZHANG, Q. et COULOIGNER, I. (2006b). An integrated approach to extracting urban road networks from high resolution multi-spectral imagery. *In 1st EARSeL Workshop of the SIG Urban Remote Sensing*, HumboldtU-niversität zu Berlin.
- [Zhao et Wang, 2010] ZHAO, L. et WANG, X. (2010). Road extraction in high resolution remote sensing images based on mathematic morphology and snake model. *3rd International Congress on Image and Signal Processing*.
- [Zhaoping, 2000] ZHAOPING, L. (2000). Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, 13(1):25–50.
- [Zhou et Troy, 2008] ZHOU, W. et TROY, A. (2008). An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing*, 29:3119–3135.