Ontologies dans les images satellitaires : interprétation sémantique des images

Samuel Andrés

▶ **To cite this version:**

Samuel Andrés. Ontologies dans les images satellitaires : interprétation sémantique des images. Autre [cs.OH]. Université Montpellier II - Sciences et Techniques du Languedoc, 2013. Français. NNT : 2013MON20059 . tel-00998692

HAL Id: tel-00998692

https://theses.hal.science/tel-00998692

Submitted on 2 Jun 2014

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.

THÈSE

Pour obtenir le grade de

Docteur

Délivré par l'**Université Montpellier 2**

Préparée au sein de l'école doctorale **Information Structures Systèmes**
Et de l'unité de recherche **UMR 228 Espace-Dev**

Spécialité: **Informatique**

Présentée par **Samuel ANDRÉS**

Les Ontologies dans les Images Satellitaires

Interprétation Sémantique des Images

Soutenue le 13 décembre 2013 devant le jury composé de

Mme Nathalie AUSSENAC-GILLES	Directeur de Recherche	IRIT	Rapporteur
M. Jean-Paul DONNAY	Professeur	Université de Liège	Rapporteur
M. Pierre GANÇARSKI	Professeur	ICube	Examinateur
Mme Anne PUISSANT	Maître de Conférences	LIVE	Examinateur
M. Olivier STRAUSS	Maître de Conférences	LIRMM	Examinateur
Mme Thérèse LIBOUREL	Professeur	Université Montpellier 2	Directeur de thèse
Mme Isabelle MOUGENOT	Maître de Conférences	Université Montpellier 2	Co-directeur de thèse

Résumé

Étant donnée l'évolution technologique des capteurs embarqués à bord des satellites, le potentiel d'images satellitaires accessible s'accroît de telle manière que se pose maintenant la question de son exploitation la plus efficace possible. C'est l'objectif du projet CARTAM-SAT que de fluidifier la chaîne de traitements depuis les satellites jusqu'aux utilisateurs des images. La thèse s'inscrit dans ce cadre.

Les traitements relatifs aux images ont évolué au cours des années. Les images basse résolution étaient traitées par une approche dite « pixel » alors que la haute résolution a permis le développement d'une approche dite « objet ». Cette dernière s'attache à analyser non plus des pixels isolés, mais des groupes de pixels représentatifs d'objets concrets sur le terrain. Ainsi, en principe, ces groupes de pixels sont dotés d'une sémantique propre au domaine de la télédétection.

La représentation des connaissances a évolué parallèlement aux images satellitaires. Les standards de représentation ont profité de l'expansion du web pour donner naissance à des standards comme OWL. Celui-ci repose en grande partie sur les logiques de description qui permettent l'utilisation de raisonneurs automatiques capables d'inférer une connaissance implicite.

Cette thèse se place à la jonction de ces deux sciences et propose une approche ontologique d'analyse des images satellitaires. Il s'agit de formaliser différents types de connaissances et de conceptualisations implicitement utilisés par les logiciels de traitement d'image et par les experts en télédétection, puis de raisonner automatiquement sur la description d'une image pour en obtenir une interprétation sémantique.

Ce principe général est susceptible de nombreuses déclinaisons techniques. La mise en œuvre a consisté en la réalisation d'un prototype alliant une bibliothèque d'analyse d'images satellitaires et un raisonneur basé sur les ontologies. L'implémentation proposée dans la thèse permet d'explorer quatre déclinaisons techniques du principe qui mènent à des discussions sur la complémentarité des paradigmes d'analyse « pixel » et « objet », la représentation de certaines relations spatiales et la place de la connaissance par rapport aux traitements.

Mots-clefs

Fossé sémantique, Images satellitaires, Ontologie, Géomatique, Télédétection, Connaissance

Summary

GIVEN THE TECHNOLOGICAL DEVELOPMENT of embedded satellite sensors, the potential of available satellite images increases so that the question now arises of their most efficient exploitation possible. This is the purpose of the CARTAM-SAT project to fluidize the processing workflow from satellite images to users. The thesis is part of this framework.

Processing operations relating to images have evolved over the years. Low-resolution images were processed by a so-called "pixel" approach while the high-resolution has allowed the development of a so-called "object" approach. The latter focuses on analysing not about the isolated pixels, but about groups of pixels representative of concrete objects on the ground. Thus, in principle, these are groups of pixels with a domain-specific remote sensing semantics.

Along with satellite imagery, knowledge representation has evolved. The standards of representation have benefited from the expansion of the web to give rise to standards like OWL. This one is widely based on description logics that allow the use of automated reasoners able to infer implicit knowledge.

This thesis is at the junction of these two sciences and provides an ontological approach for analysing satellite images. The aim is to formalize different types of knowledges and conceptualizations implicitly used by image processing programs and by remote sensing experts, and then reasoning automatically on an image description to obtain one semantic interpretation.

This general principle may have numerous technical variations. The implementation consisted in a prototype combining a satellite image analysis library and an ontology-based reasoner. The implementation proposed in the thesis allows to explore four technical variations of the principle that lead to discussions on the complementarity of "pixel" and "object" analysis paradigms, the representation of some of the spatial relations and the role of knowledge in relation to processing.

Keywords

Semantic gap, Satellite images, Ontology, Geomatics, Remote sensing, Knowledge

Sommaire

Sommaire **4**

Remerciements **7**

Préambule **9**

1 Introduction **11**
1.1 La raison du sujet 12
1.2 Le cadre institutionnel de la thèse 12
1.3 Cartographie dynamique des Territoires Amazoniens : des satellites aux acteurs . . 14
1.4 Objectifs de la thèse 14
1.5 Plan du mémoire 14

Conventions de notation **17**

I Contexte et problématique **19**

2 Généralités et spécificités des images en télédétection **21**
2.1 Introduction aux images numériques 22
2.2 La télédétection 25
2.3 Enregistrement, traitement et analyse de l'information des images de télédétection 45

3 Formalisation et conceptualisation de la connaissance **61**
3.1 Bref historique de la conceptualisation et de la représentation des connaissances . . 62
3.2 Structuration et description des données 63
3.3 Les logiques de description 67
3.4 Représentation de la connaissance et enrichissement de l'information 77

4 Les ontologies en imagerie, télédétection et données spatiales **79**
4.1 La sémantique des pixels 80
4.2 Analyse d'image orientée objet 82
4.3 Segmentation et sémantique 85
4.4 Les ontologies pour l'annotation d'images 86
4.5 Formalisation de la connaissance experte 86
4.6 Raisonnement 88
4.7 Structuration des données spatiales 89
4.8 Réflexions sur la dimension temporelle 90

II Proposition d'approche 93

5 Principe général de l'approche 95
5.1 Conceptualisation de référence 96
5.2 Connaissance contextuelle 103
5.3 Peuplement par la description de l'image 105
5.4 Interprétation sémantique de l'image 106

6 Planification de la mise en œuvre 109
6.1 Données en entrée 110
6.2 Que faire de la connaissance en sortie de la segmentation ? 113
6.3 Intégration et modularité de la connaissance 116
6.4 Réflexions sur la modularisation de la connaissance 121
6.5 Réflexions sur la représentation des relations spatiales 122

7 Déclinaisons du principe et implémentations 125
7.1 Application associée à l'analyse orientée objet 126
7.2 Intégration des relations spatiales 135
7.3 Adaptation à la description des pixels 141
7.4 Classification séparée des objets 145
7.5 Réflexions 148
7.6 L'hypothèse du monde ouvert 149

III Conclusion 155

8 Conclusion et ouverture 157
8.1 Bilan 158
8.2 Remarques relatives au déroulement des travaux 159
8.3 Perspectives 161

IV Annexes 167

A Quelques indices utilisés en traitement d'image 169
A.1 Indices spectraux 169
A.2 Indices de texture 170

B Essais d'utilisation des relations spatiales réifiées d'OBOE 171

Table des matières 175

Liste des définitions et des remarques 179

Table des figures 182

Liste des tableaux 188

Bibliographie 189

Remerciements

Mes remerciements s'adressent en premier lieu à Mme Thérèse Libourel, mon directeur de thèse et à Mme Isabelle Mougenot qui a co-dirigé ce travail. Toutes deux m'ont proposé cette thèse qui a représenté pour moi, en plus de la découverte de nouvelles disciplines, un sujet de réflexion passionnant.

Je voudrais aussi remercier Mme Nathalie Aussenac-Gilles et M. Jean-Paul Donnay, qui ont aimablement accepté d'examiner ce mémoire comme rapporteurs, ainsi que Mme Anne Puissant, M. Olivier Strauss et M. Pierre Gançarski pour leur présence dans le jury et plus largement pour l'intérêt qu'ils ont manifesté pour le sujet de thèse. Je remercie également Mme Frédérique Seyler, responsable scientifique du projet CARTAM-SAT, M. Frédéric Huynh qui m'a accueilli au sein de l'unité de recherche, ainsi que M. Jean-Christophe Desconnets. Je n'oublie pas MM. Vincent Heurtaux et Johann Sorel qui, en me donnant ma chance à *Geomatys* pour mon stage de fin d'études de Master, ont sans doute facilité cette opportunité.

L'ambiance de travail pendant ces trois années a été fort sympathique. Je pense tout particulièrement à la collaboration scientifique établie avec Damien Arvor et Marie-Angélique Laporte, encouragée par Laurent Durieux, au dialogue établi avec M. Mikhail Levin, mais également aux moments de bonne humeur partagés avec Meriam Bayoudh, Anne Toulet, Christelle Pierkot, Bertrand Guerrero, Dominique Briand, Dorian Ginane, Romain Boulet, Emmanuel Roux et toute l'équipe de *Geomatys*.

Enfin, je rends grâce au ciel de m'avoir conservé le soutien de ma famille et de mes amis en toutes circonstances. Je pense en particulier à Jean-Pierre Dugas qui a pris la peine de relire mon mémoire en détail.

À vous tous : merci.

Préambule

Ce mémoire s'articule en plusieurs sections. Tout d'abord, une introduction présente les combinaisons de champs scientifiques qui ont conduit à la formulation du sujet dans un environnement de travail original et pluridisciplinaire. Ce chapitre pose également les objectifs de la thèse, ainsi qu'un plan. Il est suivi des conventions de notation utilisées dans ce document.

Le cœur du mémoire est constitué de deux parties principales. La première présente le contexte disciplinaire dans les différents domaines concernés : l'image, la télédétection, la représentation des connaissances et les logiques de description. Elle s'achève par un tour d'horizon des propositions concernant de près ou de loin l'utilisation combinée de l'image et de la connaissance.

La seconde partie est consacrée à l'exposé de la proposition élaborée dans le cadre de la thèse. Elle contient une présentation du principe méthodologique, des discussions préalables aux applications et des propositions de mises en œuvre concrètes. Ces dernières sont commentées, illustrées dans la mesure du possible, et accompagnées de résultats avec quelques tentatives d'évaluation.

Les parties suivantes sont consacrées à la conclusion des travaux et aux annexes de documentation relatives à quelques points évoqués dans le mémoire.

1 Introduction

Avant d'entrer dans le vif du sujet, ce chapitre préliminaire présente l'environnement de la thèse et en particulier l'environnement scientifique dans lequel elle s'est déroulée. Il s'agit également de tracer les grandes lignes du projet de *Cartographie dynamique des Territoires Amazoniens : des Satellites aux Acteurs* – CARTAM-SAT qui constitue le cadre de la thèse[1]. De l'expertise de terrain à l'analyse d'image satellitaire, ce projet a été porté par une unité aux compétences scientifiques multiples.

Sommaire

1.1 La raison du sujet **12**
1.2 Le cadre institutionnel de la thèse **12**
1.2.1 L'Institut de Recherche pour le Développement 12
1.2.2 L'Espace au service du Développement 13
1.2.3 Systèmes d'Information et de Connaissance 13
1.3 Cartographie dynamique des Territoires Amazoniens : des satellites aux acteurs **14**
1.3.1 Du satellite à l'image 14
1.3.2 De l'image à la connaissance 14
1.3.3 De la connaissance aux acteurs 14
1.4 Objectifs de la thèse **14**
1.5 Plan du mémoire **14**

[1] Le cadre du sujet a été présenté par Andrés (2011)

1.1 La raison du sujet

De même que la géophysique accompagne et complète les méthodes intrusives en Sciences de la Terre, l'imagerie satellitaire apporte une information complémentaire utile à de nombreuses disciplines. En ajoutant l'observation de la Terre aux observations de terrain, elle élargit leur champ de vision. De plus, le passage d'une échelle locale à une échelle plus globale, permet d'appréhender de vastes espaces sans qu'il soit nécessaire d'arpenter chaque centimètre carré de terrain.

Les images satellitaires ont beaucoup évolué depuis leur apparition. En particulier, leur capacité à représenter des objets de plus en plus proches de l'échelle humaine s'est imposée avec l'avènement des capteurs à haute, puis à très haute résolution spatiale.

Or, qui dit « raffinement dans la représentation » dit aussi souvent « accroissement du volume de données » et « complexification de l'analyse ». Les méthodes de travail des experts en télédétection ont suivi l'évolution des images... mais c'est bien toujours la connaissance, l'expérience et l'expertise qui font la valeur ajoutée d'une analyse d'image.

Cette connaissance, conceptuelle et implicite pour les experts en télédétection, est immatérielle et difficilement accessible et manipulable en tant que telle. Pourtant, on la rencontre souvent, soit incluse dans les programmes informatiques des logiciels utilisés, soit employée plus ou moins intuitivement par les experts dans leurs analyses.

Ne pourrait-on pas utiliser les progrès en représentation des connaissances pour en prendre le contrôle et en tirer meilleur parti ?

L'énoncé « Les ontologies dans les images satellitaires » évoque clairement les principales disciplines concernées par la thèse. Ce sont en première approximation[1] :

- l'image numérique, du côté de l'information transmise par le satellite ;
- la représentation des connaissances, en ce qui concerne la perception du monde par les acteurs.

Initialement, le sujet ouvrait un large champ de recherches possibles. Restait à délimiter des contours plus précis. Ils ont été tracés de façon assez stricte au fil de la thèse, de manière à éviter le risque d'un éparpillement inhérent au travail de recherche et particulièrement accru par la pluridisciplinarité. Il est très facile en effet de passer un temps considérable exclusivement dans le domaine de l'image ou de la représentation des connaissances en télédétection. Là n'était pas l'objectif.

A priori, le sujet a donc été délimité à la jonction entre l'image et la connaissance. Il ne s'agit donc pas de s'occuper de traitement d'image ou de construction de connaissance à titre principal, mais simplement d'envisager le potentiel des ontologies pour œuvrer à leur rapprochement.

1.2 Le cadre institutionnel de la thèse

1.2.1 L'Institut de Recherche pour le Développement[2]

L'Institut de Recherche pour le Développement (IRD) est un établissement public à caractère scientifique et technologique (EPST) placé sous la double tutelle du Ministère de l'Enseignement Supérieur et de la Recherche et du Ministère des Affaires Étrangères et Européennes.

Il hérite de l'Office de la Recherche Scientifique Coloniale (ORSC). Successivement renommé Office de Recherche Scientifique d'Outre-Mer (ORSOM) puis Office de la Recherche Scientifique et Technique Outre-Mer (ORSTOM), cet organisme se consacre alors à la recherche scientifique relative à l'empire colonial, particulièrement en zone tropicale.

Après avoir accompagné l'indépendance des colonies en posant les jalons des politiques de coopération, l'ORSTOM est réorganisé en profondeur au cours des années 1980 et 1990. Il adopte son appellation actuelle d'IRD en 1998 et tisse des liens avec la recherche universitaire par le biais des Unités Mixtes de Recherche (UMR) au cours des années 2000.

[1] Le sujet prendra sa place dans l'ensemble du projet de recherche présenté en 1.3.

[2] Le résumé historique a été composé à partir de la page http://www.ird.fr/l-ird/historique [accessed : 18-02-2013]. On pourra s'y reporter pour y trouver une information plus détaillée.

1.2.2 L'Espace au service du Développement

L'Unité Mixte de Recherche 228 *L'Espace au service du Développement* (ESPACE-DEV) est l'une de ces unités dont la création a été encouragée par l'IRD. Elle est issue de l'Unité de Service Espace de l'IRD qui était chargée de la réception, du traitement et de l'exploitation des données satellitaires. La nouvelle structure de recherche a reçu le renfort d'universitaires et bénéficie des co-tutelles de l'Université Montpellier II, de l'Université Antilles-Guyane et de l'Université de la Réunion[1]. Les chercheurs, enseignants-chercheurs, ingénieurs et techniciens qui la composent, sont issus d'horizons multiples. Trois équipes de recherche de l'unité illustrent parfaitement cette jonction d'intérêts scientifiques divers.

L'équipe « Observation Spatiale de l'Environnement » (OSE) détient une grande expertise des données satellitaires, depuis le recueil de l'information jusqu'à l'analyse des systèmes et processus physiques qu'elle représente, en passant par son traitement.

L'équipe « Approche Intégrée des Milieux et des Sociétés » (AIMS) appréhende les systèmes étudiés sous un angle social et humain plutôt que par l'intermédiaire des capteurs satellitaires. Même si les questions environnementales les concernent tout autant, les experts de l'équipe AIMS ont un point de vue très différent des spécialistes de l'équipe OSE : entre l'image satellitaire et la relation humaine, l'échelle varie de même que le type d'information recueillie.

L'équipe « Systèmes d'Information et de Connaissance » (SIC) propose quant à elle une approche transversale aux points de vue que les équipes OSE et AIMS ont sur l'environnement. Les sciences de l'information allient aujourd'hui des concepts éprouvés tels les Systèmes de Gestion de Bases de Données (SGBD), à des élaborations plus récentes comme les spécifications du web sémantique. Ainsi que l'a par exemple prouvé la naissance des Systèmes d'Information Géographique, ce domaine peut être associé aux questions environnementales pour le plus grand bénéfice des deux parties[2].

Cet environnement pluridisciplinaire constitue un cadre propice à l'échange et au partage de connaissances, terrain idéal pour accueillir un travail de thèse sur la sémantique des images de télédétection (Définition 10).

1.2.3 Systèmes d'Information et de Connaissance

Par sa position médiatrice entre les deux autres équipes de recherche de l'UMR, l'équipe SIC était naturellement disposée à accueillir tout travail potentiellement concerné par la représentation des conceptualisations de l'image et de l'expertise de terrain.

Cette équipe oriente ses activités de recherche en grande partie autour des ontologies mais présente elle-même une certaine pluridisciplinarité allant du catalogage de métadonnées géographiques aux bases de données spatialisées en passant par les systèmes d'information géographique (SIG).

Définition 1 (Système d'information, d'après Reix (2004))**.** Ensemble organisé de ressources : matériel, logiciel, personnel, données, procédures... permettant d'acquérir, de traiter, de stocker, de communiquer des informations (sous formes de données, textes, images, sons...) dans et entre des organisations.

Définition 2 (Système d'information géographique)**.** Système d'information relatif à des données spatialement référencées à la surface de la Terre.

[1] L'Université d'Avignon et des Pays du Vaucluse a quant à elle signé une convention avec l'UMR.

[2] S'il est évident que les applications géographiques ou environnementales apportent un bénéfice direct à ces domaines, le retour d'expérience en informatique même n'en est pas moins intéressant. On peut citer par exemple les questions relatives à la représentation des relations spatiales (Clementini, 2009).

1.3 Cartographie dynamique des Territoires Amazoniens : des satellites aux acteurs

Le projet « Cartographie dynamique des Territoires Amazoniens : des satellites aux acteurs », financé, via l'IRD, par le Fond Européen de Développement Régional (FEDER), propose de fluidifier la chaîne d'acquisition, de traitements, d'exploitation et de diffusion des données satellitaires. Dans cette perspective, il en développe chacune des étapes en identifiant les points susceptibles d'être améliorés par des avancées techniques ou méthodologiques.

1.3.1 Du satellite à l'image

En amont de la chaîne de traitements se trouve l'acquisition des données brutes et les prétraitements. Le premier volet du projet vise à la fluidification de ces étapes, de manière à optimiser la mise à disposition des données récoltées.

1.3.2 De l'image à la connaissance

Le second volet du projet vise à rapprocher les images satellitaires de leurs utilisateurs. D'orientation plutôt méthodologique, il se décompose en deux parties.

1.3.2.1 De l'image au percept

La première partie, proche des questions de bas-niveau comme la segmentation, est fortement orientée vers l'optimisation des traitements et le développement d'algorithmes de détection de changements, de segmentation multispectrale, de recalage multisources et de désenuagement.

1.3.2.2 Du percept au concept

Le cœur de la thèse se situe dans cette seconde partie qui propose de relier les données de l'image à la connaissance de l'expert au moyen des ontologies, en examinant les aspects méthodologiques et en se dirigeant vers une solution de plate-forme opérationnelle.

1.3.3 De la connaissance aux acteurs

Enfin, la partie aval du projet est centrée autour de la problématique de la diffusion des données et des connaissances ainsi qu'autour de la formation universitaire en géomatique et télédétection.

1.4 Objectifs de la thèse

Les objectifs de la thèse consistent à examiner si les ontologies présentent un intérêt en télédétection en vue de l'interprétation et de la classification automatisées d'images satellitaires.

Dans cette perspective, la thèse s'applique en particulier à explorer le potentiel des techniques d'intelligence artificielle liées à la formalisation de la connaissance (raisonnement automatique) pour produire des méthodes de classifications que l'on espère améliorées.

C'est dans ce but que l'on envisage l'utilisation de la connaissance experte formalisée en combinaison avec les principaux paradigmes d'analyse des images actuellement utilisés en télédétection : analyse *orientée objet* ou analyse des pixels.

Enfin, il s'agit de se pencher sur les implications de cette approche originale en termes d'hypothèses préalables, de mise en œuvre et d'observation des résultats.

1.5 Plan du mémoire

La première partie du mémoire s'attache à poser les bases de la proposition. Elle débute par un chapitre consacré d'abord à l'image numérique en général qui s'oriente progressivement vers les spécificités de l'image satellitaire et pose à cette fin les notions fondamentales de la télédétection.

Le chapitre suivant présente quant à lui les langages et outils couramment utilisés en représentation des connaissances et s'attarde particulièrement sur les logiques de description et leur potentiel de raisonnement automatique. Le dernier chapitre de cette partie propose un état de l'art des travaux articulant de diverses manières les images numériques et la connaissance formalisée.

La deuxième partie concerne la proposition d'approche formulée dans le cadre de la thèse. Elle s'articule également en trois chapitres. Le premier présente le principe méthodologique. Le deuxième prépare la présentation de la mise en œuvre par des discussions autour de l'utilisation des images et de la connaissance. Enfin, le dernier chapitre de cette partie est consacré à la mise en œuvre, à la présentation des résultats et à l'évaluation de différentes applications techniques du principe de l'approche.

Enfin, une troisième partie conclut le document par un résumé accompagné de quelques réflexions et de pistes d'ouvertures qui, pour n'avoir pas été explorées au cours de la thèse, pourraient se révéler intéressantes.

Conventions de notation

On utilise dans ce mémoire des représentations graphiques à des fins d'illustrations ou d'exemples. Un effort particulier a été apporté à l'homogénéité des figures représentant les graphes sémantiques (Figure 1.1), les logiques de description (Figure 1.2) et les chaînes de traitements (Figure 1.3).

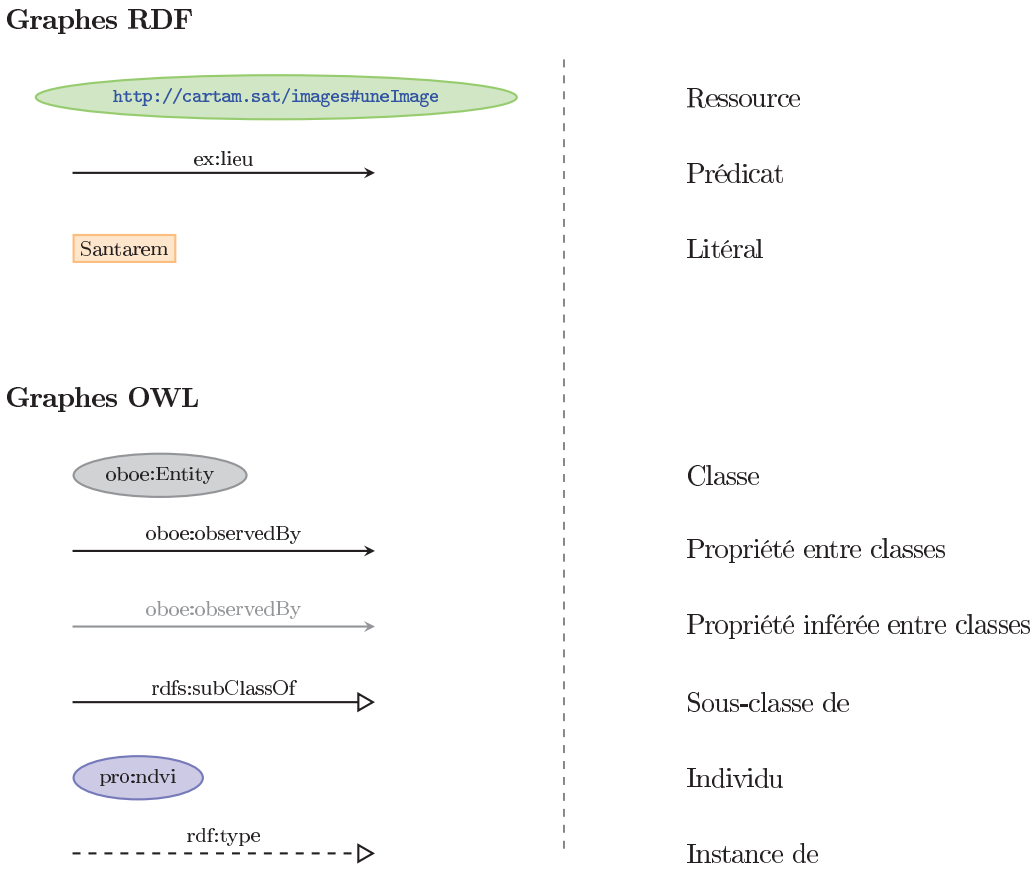
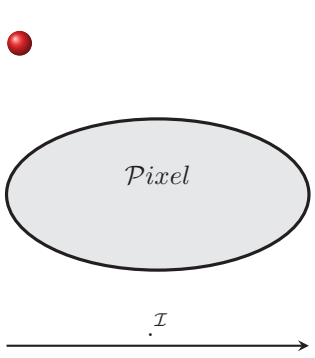


Figure 1.1 – Conventions de notation graphique.

Logiques de description



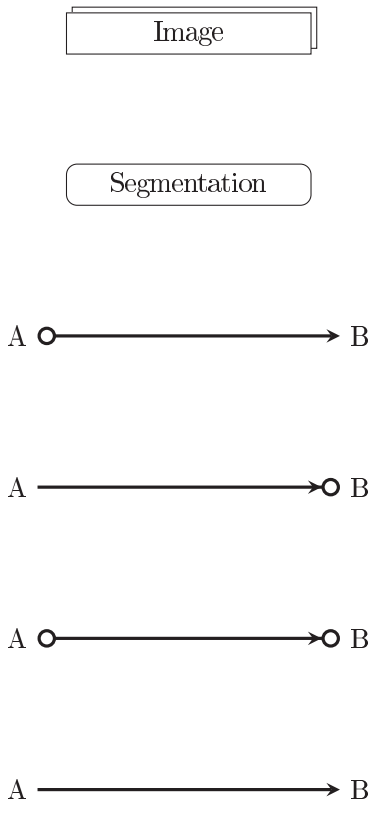
Individu

Concept

Interprétation de

FIGURE 1.2 – Conventions de notation des éléments en logiques de description.

Chaînes de traitements et descriptifs assimilés



Donnée/information/résultat

Traitement

Transition de traitement à donnée, du port A vers la donnée B

Transition de donnée à traitement, de la donnée A vers le port B

Transition de traitement à traitement, du port A vers le port B

Transition de donnée à donnée, de la donnée à vers la donnée B

FIGURE 1.3 – Conventions de notation des chaînes de traitements et descriptifs assimilés, d'après le langage formalisé par Lin (2011).

Première partie

Contexte et problématique

2

Généralités et spécificités des images en télédétection

Borné par différents domaines d'expertise tels que la télédétection et la représentation des connaissances, l'exposé des travaux de la thèse nécessite la réalisation d'un résumé suffisamment précis des domaines concernés, avant d'examiner les aspects fédérateurs de l'ensemble.

Dans cette perspective, ce chapitre propose un tour d'horizon du monde de l'image et de ses particularités en télédétection. Tout en exposant les bases de ces disciplines, il s'agit de poser les concepts et de définir les termes de façon à réduire les ambiguïtés et les malentendus.

Sommaire

2.1 Introduction aux images numériques **22**
2.1.1 L'étymologie de l'image 22
2.1.2 Généralités sur les images numériques 23
2.2 La télédétection **25**
2.2.1 Processus général 25
2.2.2 Lumière solaire et couleurs 26
2.2.3 Télédétection spatiale 28
2.2.4 Réflectance et signature spectrale 35
2.2.5 Les niveaux de prétraitement des images SPOT 43
2.2.6 Les résolutions d'une image satellitaire 44
2.2.7 La télédétection par les hyperfréquences 45
2.3 Enregistrement, traitement et analyse de l'information des images de télédétection **45**
2.3.1 Nature de l'image de télédétection 45
2.3.2 Format des images 46
2.3.3 Mise en évidence visuelle par rehaussement et filtrage, traits des images 46
2.3.4 Opérations sur les images 49
2.3.5 Notion de texture 49
2.3.6 Techniques de classification appliquées aux images 52
2.3.7 Segmentation 53

La difficulté de poser un vocabulaire est particulièrement aiguë dans un environnement pluridisciplinaire. Il est difficile, voire impossible de s'accorder pour chaque terme aux travaux des différentes communautés. Par exemple, le mot adopté de l'anglais « *feature* » n'a pas du tout le même sens selon qu'il s'agisse de traitement d'image ou de géomatique. D'autres mots, comme « objet » sont utilisés dans des domaines différents et présentent à la fois une différence de sens liée à leur contexte d'utilisation, mais également une certaine parenté sémantique commune[1].

D'ailleurs, même à l'intérieur d'une communauté, le consensus n'émerge pas toujours de manière évidente. Des mots ou des expressions peuvent engendrer des quiproquos, comme « image vectorielle ». Des choix s'imposent donc concernant le vocabulaire. Nous essayons de réaliser ces choix le plus souvent en référence à des travaux antérieurs, mais certaines définitions sont adaptées ou posées dans le cadre de la thèse.

2.1 Introduction aux images numériques

Cette introduction concerne d'une part la sémantique complexe du mot « image » ainsi que les principales caractéristiques des images numériques de manière générale.

2.1.1 L'étymologie de l'image

La polysémie du mot image en français est subtile au premier abord. Le dictionnaire de l'Académie Française recense quelques-unes des significations courantes.

- Il peut tout d'abord s'agir de la représentation d'un objet :
  - produite par la jonction de rayons lumineux ou de leur prolongement (sens physique) ;
  - réalisée par un moyen technique (dessin, sculpture, photographie).
- Sur le plan de la représentation mentale, l'image peut également signifier la ressemblance, la description ou bien une métaphore ou une idée.
- Plus accessoirement, on peut signaler le sens mathématique de l'image qui désigne la correspondance d'un élément ou d'un ensemble d'éléments par une application.

L'Institut National de Recherches Pédagogiques[2] (INRP) rappelle également que le terme recouvre indistinctement l'image comme *objet matériel* et l'image comme *représentation mentale*. En anglais cette différence apparaît dans le langage : le premier sens est désigné par le mot *picture*, le second est porté par le vocable *image*. Le recours à la conceptualisation sous-jacente à cette langue est d'ailleurs non seulement intéressant par cette distinction, mais également révélateur par le fait que pour être différentes, les deux acceptions n'en sont pas moins liées : bien que composée de pixels (*picture elements*), une image satellitaire se dit en effet *satellite image*. Il s'agirait donc à la fois d'une représentation physique et mentale. L'INRP propose respectivement les expressions d'*inscription visuo-spatiale* et de *représentation visuo-spatiale* qui permettent de lever toute ambiguïté au prix d'une lourdeur certaine. On notera au passage la présence essentielle de l'espace dans les deux expressions.

Étymologiquement, *image* dérive du latin *imitor* par *imago*, ce qui sous-entend initialement la référence à un modèle imité. Toutefois, en pratique, la source de l'image a ensuite été trouvée non plus seulement dans l'imitation de phénomènes physiques observables, mais également dans

[1] D'après l'*Open Geospatial Consortium* (http://www.opengeospatial.org/taxonomy/term/13), le terme de *feature*, « point de départ de la modélisation de l'information géographique [est] une représentation numérique d'une entité du monde réel ou une abstraction du monde réel [, c'est-à-dire,] presque tout ce qui prend place dans le temps et dans l'espace. [...] Les termes de « *feature* » et d'« objet » sont souvent utilisés comme synonymes ». Or, en imagerie, *feature* désigne plutôt une caractéristique d'une entité dans l'image.

Pour l'OGC, « Les images d'observation de la Terre sont vues comme des couvertures maillées qui contiennent des features dont les géométries sont [...] des « ensembles de pixels » considérés comme des surfaces ». « Une couverture est [elle-même] une *feature* qui associe des positions à l'intérieur d'un espace borné (son domaine spatio-temporel) à des valeurs d'attributs de *features* ».

[2] http://www.inrp.fr/Tecne/histimage/SoTeintro.htm

l'imaginaire. Par la suite, on ne s'attachera pas à ce dernier cas puisque l'image satellitaire trouve bien sa source dans des phénomènes physiques.

Le champ lexical de l'image contient un vocabulaire à la signification mouvante dans le temps. L'INRP cite par exemple les mots de *figure*, *icône* ou *forme* dont le sens varie selon l'époque et le contexte... Il s'agit d'un domaine hautement polysémique.

2.1.2 Généralités sur les images numériques

On remarquera dans le contexte informatique qu'une « image » peut prendre deux sens différents. C'est d'abord *l'information* en mémoire (*l'image comme donnée*). On pourrait faire une analogie entre ce sens et une représentation mentale. Mais c'est aussi *l'affichage* sur un support (*l'image à l'écran*). Cette seconde acception se rapporterait davantage à une inscription visuo-spatiale. Ces points de vue se discutent, mais il n'en reste pas moins que si *l'image comme donnée* représente un phénomène à un premier niveau, *l'image à l'écran* représente *l'image comme donnée* à un autre niveau. À chaque niveau, il y a distinction et transformation (voire approximation) entre le *représenté* et sa *représentation*.

La notion d'image numérique peut donc prêter à confusion. Sa nature est bien particulière car elle peut être appréhendée soit directement par les valeurs numériques qui la caractérisent, soit par le truchement de sa représentation à l'écran : une « image d'image » en quelque sorte. Selon les cas, la visualisation peut être incomplète, projetée, interpolée, calculée... et donc ne pas représenter exactement les données de l'image. C'est à ces dernières que l'on s'attache plus particulièrement ici.

2.1.2.1 Images vectorielles et images matricielles

Les images numériques recouvrent deux grandes familles bien distinctes.

L'information des *images vectorielles* représente les formes, couleurs... en s'appuyant sur des définitions géométriques (fonctions, paramètres...). L'image de l'image (affichage à l'écran...) est calculée à partir de cette représentation analytique. Les données des l'images vectorielles permettent donc des représentations graphiques ajustées à la résolution des systèmes de visualisation qui s'appuient de nos jours sur des technologies matricielles.

Il en va tout autrement des *images matricielles* qui sont des collections de valeurs numériques. Elles sont attachées à une échelle particulière, en fonction de la distance qui sépare les points de l'espace auxquels sont affectés les valeurs. Lorsque l'échelle de visualisation est plus grande, toutes les valeurs ne sont pas représentées, puisqu'elles sont plus nombreuses que le nombre de pixels de l'écran. À l'inverse, une échelle de visualisation plus faible impose d'interpoler l'affichage d'une manière ou d'une autre.

Dans une image vectorielle, les formes sont définies en quelque sorte en *intention*, alors qu'une image matricielle donne la valeur de chaque pixel de l'image ainsi définie en *extension*.

La télédétection est concernée en premier lieu par les images matricielles (également appelées *images rasters*) que sont les photographies aériennes ou les images satellitaires. Toutefois, le traitement de l'information peut amener à transformer ces données ou à les associer à des images vectorielles (ou plus simplement *images vecteurs*).

2.1.2.2 Discrétisation spatiale et discrétisation tonale

Une image matricielle est donc un ensemble discret de données. Théoriquement, les valeurs sont obtenues par échantillonnage (effectué selon un maillage) ; une fois les valeurs calculées, elles sont associées à une surface élémentaire (pavé, ou tesselle). Le pavage est indépendant du maillage à l'origine de l'échantillonnage. Dans les faits toutefois, le pavage intervient dès le calcul des valeurs (le pavage des capteurs). Dans cette configuration, chaque valeur est une synthèse du signal[1] en provenance de la part d'espace[2] correspondant au pavé qui lui est associé[3]. La surface du pavé est

[1] Intégration du signal, moyenne, etc.

[2] En télédétection, on considère que cette part d'espace est la surface dans le champ du capteur à l'instant de l'observation (appelée IFOV, pour *Instantaneous Field Of View*). http://www.nrcan.gc.ca/earth-sciences/geography-boundary/remote-sensing/fundamentals/985

[3] Dans certains cas, l'image numérique matricielle est toutefois utilisée pour illustrer des phénomènes qui ne se déroulent pas dans l'espace. Par exemple les coordonnées des résultats d'une transformation de Fourrier sont

centrée sur le point d'échantillonnage Bonnet (2010).

Définition 3 (Maillage, d'après Bonnet (2010))**.** Le maillage désigne la répartition des points d'échantillonage auxquels le signal est mesuré.

Définition 4 (Pavage, d'après Manzanera (2011); Bonnet (2010))**.** Le pavage désigne le découpage de l'espace associé aux valeurs enregistrées. Le pavage le plus courant est jointif, régulier et de taille maximale, mais ces propriétés ne sont pas obligatoires. De plus, les caractéristiques du pavage sont indépendantes du maillage.

Définition 5 (Pixel, d'après Manzanera (2011))**.** Un pixel est un élément du pavage associé à l'image numérique.

Les images numériques matricielles sont caractérisées par une double discrétisation des phénomènes qu'elles représentent. La première discrétisation est spatiale, liée à l'échantillonnage et au pavage de l'espace.

Définition 6 (Définition)**.** La définition d'une image numérique désigne son nombre de pixels. Elle peut être donnée selon chaque dimension de l'image.

Définition 7 (Résolution spatiale)**.** D'après Bonnet (2010), les définitions de la résolution spatiale sont nombreuses. Nous retiendrons ici la résolution spatiale comme le rapport du nombre de pixels de l'image à chaque dimension[1].

Par exemple, une image où un pixel représente une surface terrestre de 5 m de côté, la résolution est de 200 pixels par kilomètre.

Comme la discrétisation spatiale, la seconde discrétisation est intrinsèquement liée à la nature du monde dans lequel les données sont représentées. Les enregistrements numériques ne sont pas en mesure de représenter exactement une infinité de valeurs. Cette étape est également appelée *quantification* (Manzanera, 2011).

Définition 8 (Résolution tonale, d'après Manzanera (2011))**.** La résolution tonale d'une image désigne le nombre de valeurs possibles pour un pixel[2]

Remarque 1 (Niveaux de gris, d'après Manzanera (2011))*.* Par allusion à la manière courante de les visualiser à l'écran sur une échelle de gris allant du noir pour la valeur la plus basse au blanc pour la valeur la plus élevée, les valeurs sont parfois désignées comme *niveaux de gris*.

fréquentielles. On se placera généralement dans le cas d'images dont les valeurs sont associées à des coordonnées spatiales.

[1]Voir http://netia59a.ac-lille.fr/~cfrsdunkerque/article.php3?id_article=44 [accessed : 30-01-2013].

[2]L'extension de cette notion d'imagerie numérique à la télédétection n'est pas très pertinente. Dans les images multispectrales, un pixel est en effet associé à plusieurs valeurs, enregistrées chacune sur une plage indépendante.

Définition 9 (Dynamique d'image, d'après Manzanera (2011))**.** La dynamique de l'image désigne le nombre de bits utilisé pour le codage des valeurs des pixels.

Corollaire 1 (Dynamique d'image (Manzanera, 2011))**.** *Soit N la résolution tonale de l'image. On calcule la dynamique D par :* $D = log_2(N)$

Dans le cas le plus commun des photographies numériques, les valeurs sont indirectement représentatives d'une énergie électromagnétique. C'est également le cas pour les images satellitaires. Mesurées par des capteurs imageurs ou non-imageurs, d'autres grandeurs physiques peuvent être à l'origine d'images numériques comme l'énergie acoustique, gravitationnelle... En géophysique par exemple, on peut citer les profils sismiques[1].

2.2 La télédétection

Les images satellitaires sont des images numériques matricielles qui présentent certaines spécificités, notamment par rapport aux photographies numériques. Cette section expose les principes de base de la télédétection et les différents processus techniques ou naturels qui exercent une influence sur les images satellitaires.

Définition 10 (Télédétection)**.** La télédétection est la technique qui, par l'acquisition d'images, permet d'obtenir de l'information sur la surface de la Terre sans contact direct avec celle-ci.

La télédétection englobe tout le processus qui consiste à capter et à enregistrer l'énergie d'un rayonnement électromagnétique émis ou réfléchi, à traiter et à analyser l'information, pour ensuite mettre en application cette information.[2]

Définition 11 (Photointerprétation)**.** Elle consiste à étudier, analyser et interpréter des photographies aériennes ou des images satellitaires.

Remarque 2 (Télédétection et photointerprétation)**.** Bien que les deux termes sont couramment employés l'un pour l'autre, on peut différencier la télédétection de la photointerprétation. La télédétection qui s'emploie à *détecter à distance*, en particulier par l'usage d'images satellitaires, ne se limite pourtant pas à la photointerprétation si on considère qu'elle inclut potentiellement beaucoup d'autres formes de données (relevés de pluviométrie, mesures de salinité, modèle numérique de terrain, cartes géologiques...).

2.2.1 Processus général

Pouget (2005) décompose le processus général de la télédétection de la manière suivante :

[1] Le cas de la géophysique en général et des profils sismiques en particulier est intéressant. Comme pour les images satellitaires, les valeurs brutes de temps doivent être converties en mesures représentatives de profondeurs (ce qui n'est pas trivial).

[2] Définition relevée sur le site de *Ressources Naturelles Cananda* http://www.rncan.gc.ca/sciences-terre/limite-geographique/teledetection/11692 [accessed : 14-02-2013].

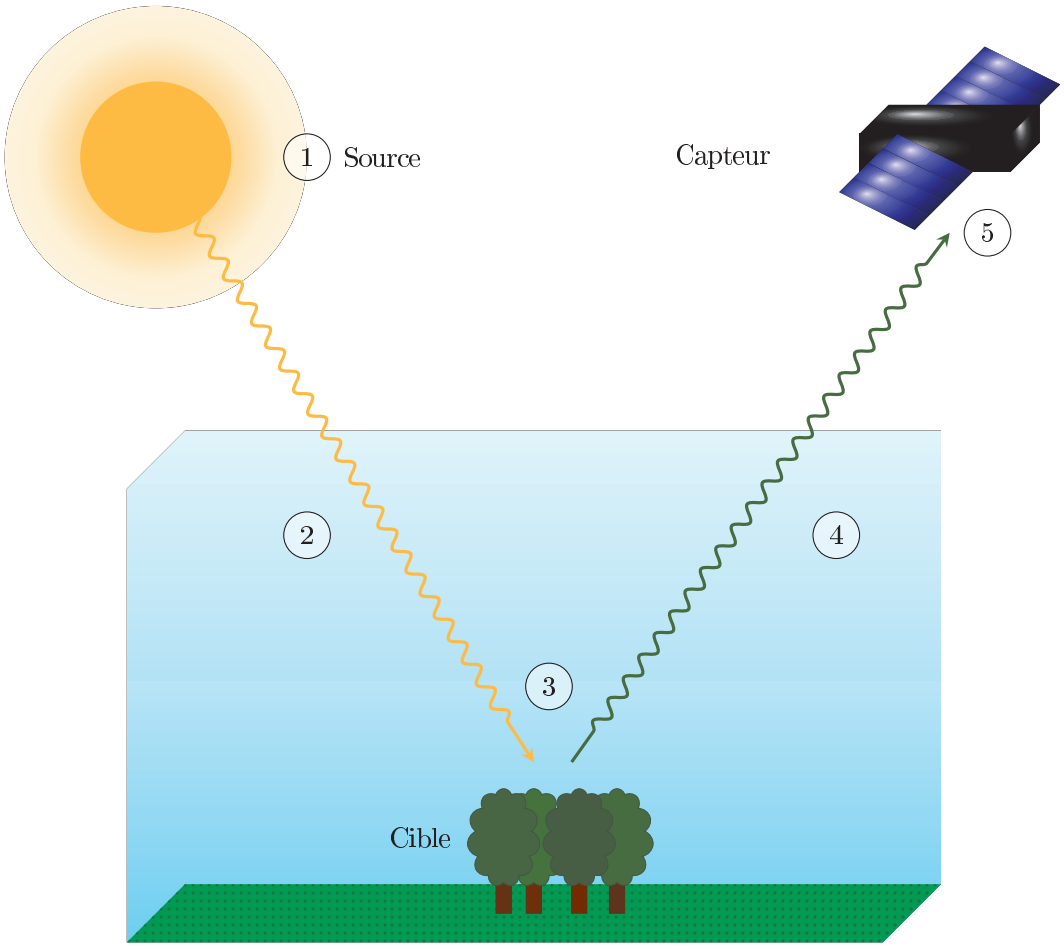


FIGURE 2.1 – Principe de la télédétection.

1. une source d'illumination émet l'énergie nécessaire à la télédétection des objets (soleil, RADAR...) ;
2. l'énergie traverse l'atmosphère jusqu'à la cible ;
3. l'énergie interagit avec la cible ;
4. une partie de l'énergie résultant de l'interaction traverse l'atmosphère jusqu'au capteur ;
5. l'énergie est recueillie et enregistrée par le capteur après discrétisation ;
6. l'enregistrement est transmis à une station de réception ;
7. l'image numérique de la cible est interprétée et analysée visuellement et/ou numériquement ;
8. enfin, l'information produite est utilisée pour mieux comprendre la cible *elle-même*.

Les cinq premières étapes de ce processus sont représentées en FIGURE 2.1. En l'absence de source d'illumination extérieure et lorsque la cible émet son propre rayonnement, les trois premières étapes n'interviennent pas dans le processus.

2.2.2 Lumière solaire et couleurs

La télédétection de la plupart des objets dépend d'une source d'énergie incidente, souvent le Soleil. L'aspect de leur image dépend d'une part de leur *nature* propre et d'autre part de la *lumière*

incidente. De plus, il peut varier au cours du temps (par exemple sous l'effet des saisons). Enfin, certains objets émettent leur propre lumière.

2.2.2.1 Notion de couleur

La notion de couleur est abstraite (Pouget, 2005) et produite par la perception des différentes longueurs d'ondes lumineuses par notre cerveau (de 380 à 760 nanomètres – soit 10^{-9} mètres).

Plusieurs systèmes permettent de reconstituer l'ensemble des couleurs visibles à partir d'une base de trois paramètres. Le plus connu est le système RVB[1] (Rouge-Vert-Bleu). Ce système, de « synthèse additive », est adapté à la représentation des couleurs des lumières. Il s'est répandu du fait de son utilisation par les écrans ou les capteurs photographiques. D'autres systèmes, de « synthèse négative », sont plus adaptés à la représentation des couleurs des objets. C'est le cas du CMJ (Cyan-Magenta-Jaune) qui sert de base aux systèmes d'impression qui ne mélangent pas des lumières, mais des encres. Enfin, des systèmes comme le TSV (Teinte-Saturation-Valeur) essayent de décrire les couleurs en se basant sur une caractérisation plus familière à la perception humaine.

De manière plus générale, nous retiendrons la Définition 12, qui bien que discutable et qualitative, présente l'avantage de ne pas se restreindre aux longueurs d'ondes visibles.

Définition 12 (Couleur, d'après Bonnet (2005))**.** La couleur d'un point d'une image est définie par une fonction de densité spectrale qui caractérise la densité énergétique de chaque longueur d'onde en ce point.

2.2.2.2 Le spectre électromagnétique

Bien que cette classification soit arbitraire, les différentes longueurs d'ondes du spectre solaire sont traditionnellement classées de la manière suivante, pour une longueur d'onde λ :

- rayons cosmiques ($\lambda < 1$ pm) ;
- rayons gamma (1 pm $< \lambda < 10$ pm) ;
- rayons X (10 pm $< \lambda < 10$ nm) ;
- ultraviolet (10 nm $< \lambda < 380$ nm) ;
- visible (380 nm $< \lambda < 780$ nm) ;
- infrarouge (780 nm $< \lambda < 1$ mm) ;
- ondes radio ($\lambda > 1$ mm).

Cette classification est un modèle qualitatif qui indique les écarts de longueurs d'ondes les plus couramment désignés. On peut également s'intéresser à la désignation qualitative des longueurs d'ondes dans le visible pour lesquelles ont peut aussi utiliser une classification :

- Violet (380 nm $< \lambda < 435$ nm) ;
- Indigo (435 nm $< \lambda < 465$ nm) ;
- Bleu (465 nm $< \lambda < 482$ nm) ;
- Cyan (482 nm $< \lambda < 492$ nm) ;
- Vert (492 nm $< \lambda < 560$ nm) ;
- Jaune (560 nm $< \lambda < 590$ nm) ;
- Orange (590 nm $< \lambda < 620$ nm) ;
- Rouge (620 nm $< \lambda < 780$ nm) ;

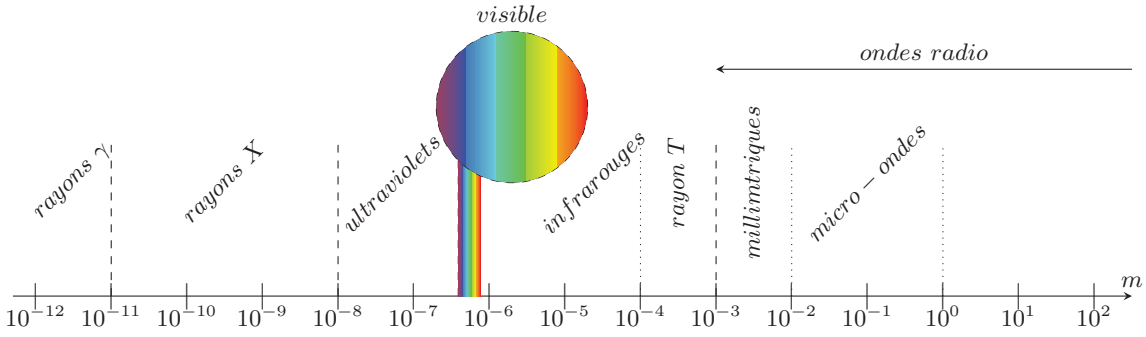


FIGURE 2.2 – Le spectre électromagnétique.

Les longueurs d'ondes accessibles à l'œil humain sont donc très limitées (FIGURE 2.2).

Or, comme la lumière visible, les longueurs d'ondes invisibles qui atteignent les objets fournissent en retour une information dépendant en partie de la nature de la surface qui les a réfléchies. Cette information inaccessible à l'œil peut toutefois être recueillie par des capteurs artificiels.

2.2.3 Télédétection spatiale

Dans le cas particulier des images satellitaires, les données sont recueillies par un capteur embarqué à bord d'un satellite d'observation de la Terre. Nous allons donc rapidement présenter ce domaine, ses concepts de base et son vocabulaire.

On s'intéresse ici aux données sous forme de raster[1], un cas particulier d'image matricielle.

Définition 13 (Image monospectrale, canal)**.** Nous appellerons *image monospectrale* toute image de télédétection en « niveaux de gris », c'est-à-dire dont les pixels sont chacun associés à une unique valeur scalaire. Une image monospectrale est accessible via un *canal* qui peut être caractérisé par une longueur d'onde ou un intervalle de longueurs d'onde.

Les filtres parfaits n'existent pas. Les acquisitions d'images monospectrales intègrent donc un signal sur une certaine fenêtre spectrale. Lorsqu'on s'intéresse à une longueur d'onde précise, on utilise des fenêtres caractéristiques relativement étroites (images spectrales). Dans d'autres cas, par exemple lorsqu'on souhaite privilégier la résolution spatiale, la fenêtre est plus large (images panchromatiques).

Définition 14 (Image spectrale)**.** On appellera ici *image spectrale* une image monospectrale dont la fenêtre spectrale est suffisamment « étroite » pour être assimilée à une longueur d'onde caractéristique.

[1] Également désigné par les initiales anglaises *RGB* (*Red-Green-Blue*).

[1] D'après l'*Open Geospatial Consortium* (http://www.opengeospatial.org/taxonomy/term/13), historiquement, ce terme provient de l'infographie. « Un raster est représentation de données spatiales en une matrice de cellules valuées. [...] En traitement de données spatialisées, un raster est la représentation de l'étendue d'un ensemble de données géographiques dans les cellules structurées en grille d'une matrice. Un affichage raster construit une image à partir de pixels, petits éléments d'image carrés, de résolution plus ou moins fine. Une base de données raster maintient une « image » de la réalité dans laquelle chaque cellule enregistre une sorte d'information moyennée sur la surface de la cellule. »

Définition 15 (Image multispectrale)**.** Une *image multispectrale* est une collection de plusieurs images monospectrales de la même scène, généralement prises depuis la même plateforme (voire le même capteur) dans un intervalle de temps très court. Chacune de ces composantes est enregistrée via son propre *canal*.

Ainsi, dans une image multispectrale, une scène peut être observée dans des fenêtres de longueurs d'ondes différentes. Chaque pixel n'est plus associé à une unique valeur scalaire mais à plusieurs, relatives chacune à une fenêtre spectrale donnée. On peut également considérer que les valeurs d'une image monospectrale sont scalaires et les valeurs d'une image multispectrale des vecteurs[1].

L'image panchromatique est un cas particulier d'image monospectrale.

Définition 16 (Image panchromatique)**.** Par opposition à l'image spectrale, l'image panchromatique est une image monospectrale dont la fenêtre spectrale est « large ». Les longueurs d'ondes concernées s'étendent souvent sur l'ensemble du spectre visible.

Certaines notions qualitatives essentielles font leur apparition : pour les couleurs d'abord, pour la largeur des fenêtres spectrales maintenant. Qu'est-ce qui permet de considérer qu'une fenêtre spectrale est étroite ou large ? Il n'y a pas de valeur déterminée... et pourtant, l'usage et l'expérience ne laissent aucun doute sur la différence entre une image spectrale est une image panchromatique. Nous essayerons donc de définir les termes de manière à réduire les ambiguïtés sans pour autant éliminer la richesse du qualitatif lorsqu'elle apparaît essentielle. En l'occurrence, poser arbitrairement une borne quantitative qui permettrait de distinguer une fenêtre large d'une fenêtre étroite n'aurait pas beaucoup de sens en télédétection[2]. En effet, la largeur d'une fenêtre spectrale est relative aux capteurs et aux longueurs d'ondes considérées.

Remarque 3 (Résolution spatiale en panchromatique et en multispectral)*.* En général, pour les capteurs d'un satellite donné, la résolution spatiale des images panchromatiques est supérieure à la résolution spatiale des images multispectrales (Voir par exemple la TABLE 2.1 page 31 pour Landsat-7 et la TABLE 2.2 page 32 pour Spot-5).

Remarque 4 (Expression courante de la résolution spatiale)*.* La résolution spatiale de la Définition 7 page 24 est habituellement inversée. Ainsi, une image sur laquelle un pixel représente une surface terrestre de 5 m de côté, la résolution est de 200 pixels par kilomètre. Mais on parlera couramment d'une « image à 5 m de résolution ».

Par rapport à la définition retenue pour la résolution spatiale concernant d'autres catégories d'images matricielles, ce raccourci de langage permet de se référer à la surface du terrain représenté par l'image, et non pas à la taille de l'image elle-même. Cela montre déjà toute l'ambiguïté qui peut exister quant à la place très intrusive du terrain dans ses représentations.

La valeur des pixels d'une image multispectrale dans un canal donné peut être visualisée en niveaux de gris. Les couleurs sont obtenues en associant les données de différents canaux aux faisceaux rouge, vert et bleu du moniteur. Cette technique de *représentation en fausses couleurs* permet de synthétiser les informations de trois canaux différents.

[1] Le terme *vecteur* ne fait pas ici référence aux images vectorielles, mais à la structure des valeurs des pixels de l'image matricielle. Ce vocabulaire peut prêter à confusion, néanmoins, *dans le contexte des images satellitaires*, il est courant de désigner les images multispectrales par le terme d'images-vecteurs.

[2] Cette notion de « largeur » dépend aussi de la position de la fenêtre dans le spectre.

2.2.3.1 Caractéristiques orbitales des satellites

Les satellites *géostationnaires* sont en orbite à environ 36000 km de la Terre et observent toujours la même zone de la surface terrestre.

L'orbite des satellites *quasi-polaires* passe au voisinage de l'axe nord-sud de la Terre. Souvent, ils ont également une orbite *héliosynchrone*, c'est-à-dire qu'ils observent toujours un endroit à la même heure locale[1]. Cette caractéristique est importante pour les observations d'une région à une saison donnée sur plusieurs années, ou bien sur plusieurs jours au sein d'une même saison, car elle garantit des conditions d'illumination similaires.

On désigne par le terme de *fauchée* la zone de la Terre observée par le satellite. Le *nadir* est l'ensemble des points qui se trouvent exactement au-dessous de la trajectoire du satellite.

2.2.3.2 Différents satellites d'observation de la Terre et capteurs embarqués

Les satellites météorologiques

Les débuts En 1960, les États-Unis d'Amérique lancèrent le premier satellite météorologique, TIROS-1, lequel fut suivi par d'autres satellites à orbite polaire héliosynchrone.

En 1966, la NASA lança le satellite géostationnaire ATS-1 (*Applications Technology Satellites*) qui produisit des images hémisphériques de la Terre toutes les 30 minutes. La série ATS fut utilisée à des fins d'observation météorologique, bien que ce ne fut pas là son objet initial.

Généralement, les satellites météorologiques ont une résolution spatiale grossière ($\approx$ 1 km) et une résolution temporelle élevée.

Définition 17 (Résolution temporelle)**.** [On définit ici la *résolution temporelle* par] le temps que prend un satellite pour [...] observer de nouveau exactement la même scène à partir du même point dans l'espace.[2]

Remarque 5 (Résolution temporelle)**.** Pour la plupart des satellites, la résolution temporelle dépend avant tout de l'orbite qui définit la fréquence de passage de la plate-forme au-dessus du même point de la surface terrestre.

Dans le cas des satellites géostationnaires, la résolution temporelle est double : nulle en ce qui concerne les points hors de son champ d'observation et théoriquement infinie en ce qui concerne les points situés à l'intérieur. La résolution temporelle s'entend alors de la période d'acquisition des images.

***Geostationary Operational Environmental Satellite* (GOES) et METEOSAT** Les satellites GOES font partie d'un réseau mondial de satellites alignés autour de la Terre à 70° de longitude d'écart. Les deux satellites géostationnaires GOES succèdent à la série ATS et fournissent des images fréquentes et à petite échelle de la surface de la Terre.

Les satellites GOES observent différents phénomènes (nuages, brumes, atmosphère, vents, température de surface des océans...) en s'appuyant sur 5 bandes spectrales.

Définition 18 (Bande spectrale)**.** Une *bande spectrale* est un intervalle (ou fenêtre) du spectre électromagnétique généralement suffisamment étroit pour être assimilé à une unique longueur d'onde.

[1] L'angle entre le plan orbital du satellite et le rayon Terre-Soleil est constant.

[2] Définition construite à partir du site de *Ressources Naturelles Cananda* http://www.rncan.gc.ca/sciences-terre/limite-geographique/teledetection/fondements/1633 [accessed : 14-02-2013].

Remarque 6 (Bande spectrale, canal). Dans le contexte de la télédétection, on emploie souvent par extension les expressions *bande spectrale* et *canal* pour désigner non pas l'intervalle du spectre, mais les données (images spectrales) qui ont été enregistrées dans l'intervalle de longueurs d'ondes correspondant.

Côté européen, les satellites METEOSAT ont une orbite géostationnaire et prennent des images à la même fréquence que les satellites du réseau GOES.

Les satellites d'observation de la Terre

Les satellites LANDSAT Les satellites Landsat sont en orbite polaire héliosynchrone.

Landsat-1 a été lancé par la NASA en 1972. Jusqu'à Landsat-4, les satellites sont équipés du capteur MSS (*Multi Spectral Scanner*) à quatre bandes, de résolution spatiale de 80×80 m^2. À partir de 1982, Landsat-4 s'est vu équipé en plus par le capteur TM (*Thematic Mapper*) qui dispose de sept bandes spectrales plus étroites, et améliore la résolution spatiale. À partir de Landsat-6, le capteur amélioré ETMP (*Enhanced Thematic Mapper Plus*) a remplacé le capteur TM... mais Landsat-6 a été perdu. Landsat-7, lancé en 1999, dispose ainsi d'un huitième canal[1] (panchromatique), mais subit des dysfonctionnements. Landsat-5 est donc le dernier satellite Landsat encore pleinement opérationnel.

La NASA prépare la relève de ces satellites avec le programme *Landsat Data Continuity Mission* (LDCM).

Bande spectrale	Fenêtre spectrale	Résolution spatiale
1	$0,45 - 0,52$ µm	30×30 m^2
2	$0,53 - 0,61$ µm	30×30 m^2
3	$0,63 - 0,69$ µm	30×30 m^2
4	$0,78 - 0,90$ µm	30×30 m^2
5	$1,55 - 1,75$ µm	30×30 m^2
6	$2,09 - 2,35$ µm	30×30 m^2
7	$10,4 - 12,5$ µm	60×60 m^2
Panchromatique	$0,52 - 0,9$ µm	15×15 m^2

Table 2.1 – Les bandes spectrales des images Landsat 7 (ETMP).

Système Pour l'Observation de la Terre (SPOT) Le système a été conçu par le CNES de la France en coopération avec la Belgique et la Suède. Les images sont commercialisées par la société SpotImage.

SPOT-1 a été lancé en 1986. Les lancements ont lieu tous les trois ou quatre ans sur une orbite polaire héliosynchrone à 830 km^2.

SPOT-1, SPOT-2 et SPOT-3 ont chacun deux capteurs HRV (Haute Résolution Visible) qui peuvent fonctionner sous deux modes :

- multispectral à 3 bandes avec une résolution de 20 mètres ;
- panchromatique avec une résolution de 10 mètres.

L'angle de visée des capteurs peut varier de 27 ° de chaque côté par rapport à la verticale. Ce *dépointage* augmente les chances d'acquisition de scènes sans nuages et permet une vision stéréoscopique.

[1] http://www.gis.usu.edu/~doug/RS5750/resources/sensors.html [accessed : 08-08-2013].
[2] La répétitivité est de 26 jours.

SPOT-4 lancé en 1998 dispose d'un canal supplémentaire dans le moyen infrarouge grâce aux capteurs HRVIR (Haute Résolution Visible et Infra-Rouge) qui succèdent aux HRV.

SPOT-5 lancé en 2002 remplace les HRVIR par deux HRG (Haute Résolution Géométrique) qui apportent une grande amélioration de la résolution spatiale ($2,5$ m à 5 m en panchromatique et 10 m en multibandes). Un instrument supplémentaire HRS (Haute Résolution Stéréoscopique) fonctionne en mode panchromatique et permet une acquisition stéréoscopique.

Définition 19 (Photogrammétrie)**.** Technique de positionnement dans l'espace des points d'une image aérienne ou satellitaire, en utilisant une autre image prise d'un point de vue légèrement différent, à la manière de la vision stéréoscopique.

Bande spectrale	Fenêtre spectrale	Résolution spatiale
1	$0,5 - 0,59$ µm	10×10 m^2
2	$0,61 - 0,68$ µm	10×10 m^2
3	$0,78 - 0,89$ µm	10×10 m^2
4	$1,58 - 1,75$ µm	20×20 m^2
Panchromatique	$0,48 - 0,71$ µm	5×5 m^2

Table 2.2 – Les bandes spectrales des images Spot 5 (HRG).

Surveillance de l'Environnement Amazonien par Satellite (SEAS) La plate-forme technologique SEAS, lancée par l'IRD et inaugurée en 2006[1] est un pôle de recherche et de formation reconnu sur le plan technologique et thématique.

SEAS appuie le Pôle Universitaire Guyanais dans ses thématiques de recherche, et particulièrement celles relatives à l'aspect « télédétection » de la dynamique et gestion des territoires amazoniens. Les activités de recherche sur la connaissance des milieux et des sociétés sont également concernées, avec la collaboration de nombreux organismes de recherche.

Sur le plan de la formation universitaire, il s'agit de mettre en place un cadre permettant d'acquérir une spécialisation tropicale confortée par des activités liées aux données issues des satellites d'observation de la Terre.

Une antenne de réception recueille en temps réel les images envoyées depuis les satellites Spot 2, 4 et 5. Les données sont ensuite traitées par le centre IRD de Cayenne.

Cette plate-forme est concernée au premier chef par le projet CARTAM-SAT.

2.2.3.3 Les systèmes de balayage

Les capteurs électroniques des satellites d'observation ont un champ de vision étroit et utilisent un système de balayage de manière à couvrir de grandes surfaces. On utilise principalement deux systèmes[2] (Pouget, 2005) :

Le balayage perpendiculaire à la trajectoire s'effectue d'un côté du capteur à l'autre en utilisant un miroir rotatif (Figure 2.3).

Le balayage parallèle à la trajectoire remplace le miroir par un ensemble de capteurs alignés : le balayeur à barrettes[3] (Figure 2.4).

Il faut donc bien comprendre que les images satellitaires ne sont pas des photographies mais des enregistrements séquentiels. Leur présentation sous forme d'image matricielle peut être trompeuse

[1] Voir http://www.ird.fr/toute-l-actualite/actualites/communiques-et-dossiers-de-presse/inauguration-de-la-plate-forme-technologique-seas-guyane#chapitre2 [accessed : 14-02-2013].

[2] Voir les illustrations proposées http://eduscol.education.fr/orbito/system/capteur/capt3.htm [accessed : 14-02-2013].

[3] C'est le cas des satellites SPOT.

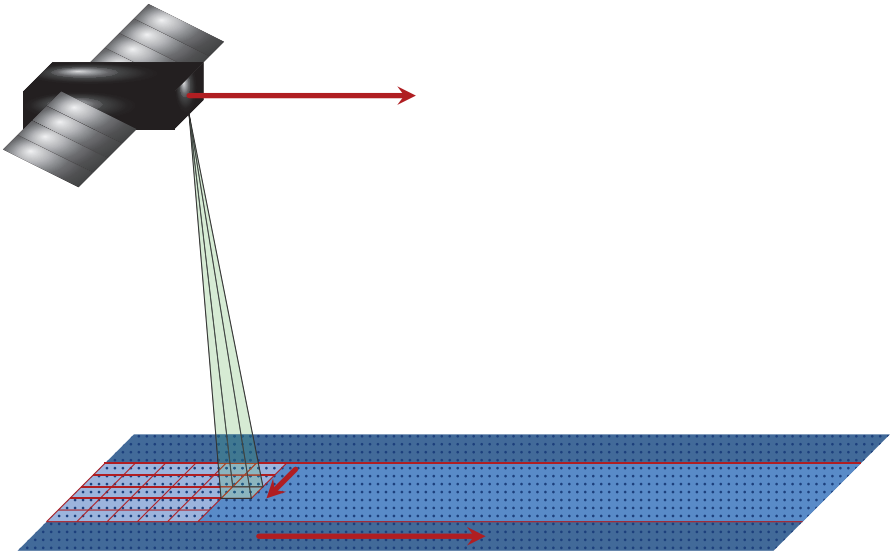


FIGURE 2.3 – Le système de balayage perpendiculaire (à miroir rotatif). Dans chaque ligne, chaque pixel est enregistré successivement, au fur et à mesure de la rotation du miroir associé au capteur. Les lignes se succèdent au fur et à mesure du déplacement du satellite. Avec ce type de capteur, les pixels d'une même ligne ne sont pas enregistrés au même instant et ils correspondent donc chacun à une position différente du satellite.

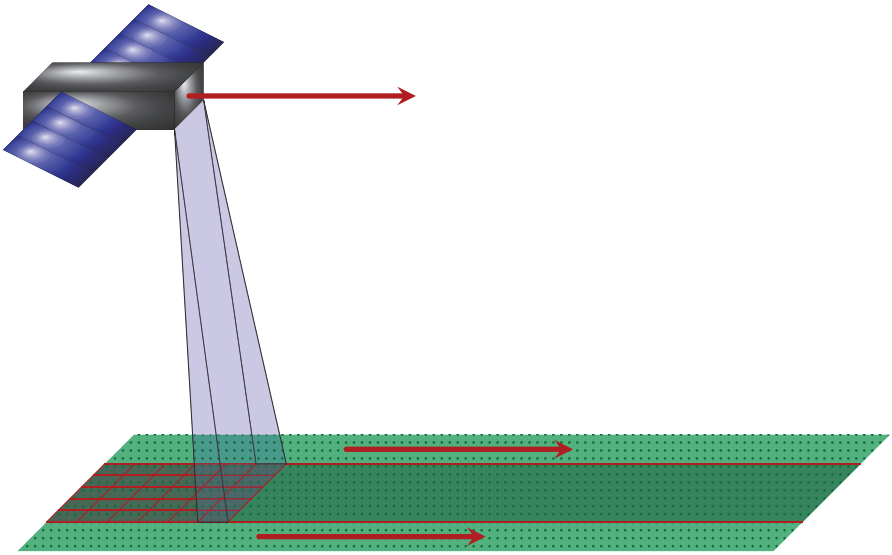


FIGURE 2.4 – Le système de balayage parallèle (à barrette de capteurs). Dans chaque ligne, tous les pixels sont enregistrés simultanément. Les lignes se succèdent au fur et à mesure du déplacement du satellite. Avec ce type de capteur, les pixels d'une même ligne sont enregistrés au même instant et correspondent donc à une unique position du satellite.

et faire croire à l'acquisition simultanée des pixels. Or, tous les enregistrements n'ont pas été effectués au même instant.

Le satellite, mobile par rapport à la cible, entraîne le capteur dans son mouvement. Cela signifie que deux données enregistrées à un moment différent ont *a priori* été mesurées depuis des endroits différents.

2.2.3.4 Les distorsions géométriques

Le mécanisme de balayage et le mouvement du satellite suggèrent que l'image satellitaire est une représentation de la réalité qui comporte un certain nombre de *distorsions géométriques*. En réalité, celles-ci sont provoquées par de nombreux facteurs :

- effet de perspective de l'optique du capteur,
- mouvement du système de balayage,
- mouvement et stabilité de la plate-forme,
- altitude, vitesse et comportement de la plate-forme,
- relief à la surface,
- courbure et rotation de la Terre.

Ces effets doivent être corrigés si l'on souhaite que chaque donnée enregistrée puisse être associée à des coordonnées géographiques.

2.2.3.5 Calibration géométrique

Les corrections géométriques consistent souvent à réaliser les étapes suivantes (Pouget, 2005) :

1. identifier des points de l'image pour déterminer leur véritable position en coordonnées au sol (points de contrôle au sol) ;
2. les résultats sont traités par ordinateur et produisent des équations de transformation qui donnent les nouvelles coordonnées au sol à partir des coordonnées originales (ligne, colonne).
3. le processus de correction géométrique aboutit à une nouvelle image avec une nouvelle géométrie.

La valeur des pixels sur l'image de sortie est obtenue par rééchantillonage suivant trois méthodes principales (Pouget, 2005) :

- le « plus proche voisin » qui utilise la valeur du pixel de l'image originale le plus proche de la localisation du nouveau pixel dans l'image.
- « l'interpolation bilinéaire »prend une moyenne pondérée par la distance des quatre pixels de l'image originale les plus proches.
- « l'interpolation cubique » se base elle sur les 16 pixels de l'image originale les plus proches.

Le *plus proche voisin* est la méthode la plus simple et n'altère pas la valeur originale, mais elle duplique certaines valeurs et en perd d'autres. Les autres méthodes produisent des valeurs *nouvelles* qui ont pour effet d'adoucir les images.

Définition 20 (Orthorectification, ou redressement)**.** Par l'*orthorectification* d'une image, les distorsions sont supprimées géométriquement, en créant une image planimétrique à chaque localisation avec une échelle cohérente dans toutes les parties de l'image[1].

Cette étape de correction prend en considération les effets de déformations géométriques dus au relief, grâce à un modèle numérique de terrain qui représente la surface orographique.

D'après Sanz et al. (2008), le redressement consiste à rectifier une « image oblique » pour obtenir en sortie une « image verticale » corrigée de toutes ou la plupart des déformations inhérentes à la prise de vue (angle, mouvement de plateforme type roulis, etc.) et aux distorsions causées par l'environnement (courbure de la Terre dans le cas des images à petites échelles, variation d'altitude au sol, etc.). Les corrections géométriques incluraient quant à elles le géoréférencement.

[1] D'après http://help.arcgis.com/fr/arcgisdesktop/10.0/help/index.html#//009t000000ms000000 [accessed :18-02-2013].

Définition 21 (Géoréférencement)**.** Le géoréférencement consiste à relier les données de l'image redressée à un système de coordonnées géographique ou cartographique.

Le géoréférencement de l'image satellitaire permet de localiser l'origine du signal à la surface terrestre.

2.2.4 Réflectance et signature spectrale

Il s'agit ici d'expliciter le lien que l'on peut établir entre l'information numérique de l'image et la nature des objets desquels provient l'énergie qui atteint le capteur d'un satellite.

2.2.4.1 Luminance et éclairement

Les valeurs de l'image satellitaire brute sont de simples *comptes numériques*, c'est-à-dire qu'elles ne sont exprimées dans aucune unité qui leur donnerait un sens physique. De fait souvent codées sur un octet, il s'agit de valeurs comprises entre 0 et 255. Elles ne peuvent être converties en une grandeur physique, qu'en tenant compte des métadonnées relatives aux capteurs jointes à l'image.

Au sens large, une métadonnée est une information à propos d'une autre information.

Définition 22 (Métadonnée)**.** Une *métadonnée* est une donnée sur une autre donnée à propos de laquelle elle fournit en général une information contextuelle (date, lieu d'acquisition ou autre...).

L'usage des métadonnées remonte en pratique à la nuit des temps. Le fait par exemple de lire le titre d'un livre sur la tranche dans une bibliothèque de manière à trouver plus facilement celui que l'on cherche, revient à utiliser des métadonnées.

Avec le développement des sciences de l'information, les métadonnées ont acquis un rôle souvent fondamental. C'est le cas par exemple dans les systèmes de gestion de bases de données relationnelles qui utilisent des tables spécifiques pour représenter les schémas des bases de données : ces *méta-tables* contiennent des informations relatives aux données (et à leur structuration), donc des métadonnées.

Des définitions plus spécifiques existent en fonction du domaine d'application.

Remarque 7 (Métadonnée, d'après le *World Wide Web Consortium*). Une métadonnée est une information compréhensible par une machine portant sur une ressource web ou d'une autre nature[1].

Remarque 8 (Métadonnée géographique, d'après la présentation de la norme ISO 19115:2003). La norme ISO 19115:2003 définit le schéma requis par la description de l'information et des services géographiques. Il fournit une information sur l'identification, l'étendue, la qualité, la convention spatiale et temporelle, la référence spatiale et la distribution des données géographiques numériques. [...] Bien que la norme ISO 19115:2003 soit applicable aux données numériques, ses principes peuvent être étendus à beaucoup d'autres formes de données géographiques telles que les cartes, les représentations graphiques, les documents textuels, aussi bien qu'à des données non-géographiques[2].

[1] http://www.w3.org/DesignIssues/Metadata.html [accessed : 01-02-2013].

```
<?xml version='1.0' encoding='UTF-8'?>
<?xml-stylesheet href="STYLE.XSL" type="text/xsl" ?>

<Dimap_Document xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
 xsi:noNamespaceSchemaLocation='Spot_Scene.xsd' name='METADATA.DIM'>


  <!-- /// Informations sur le fichier. /// -->
  <Metadata_Id>
    <METADATA_FORMAT version='1.1'>DIMAP</METADATA_FORMAT>
    <METADATA_PROFILE>SPOTSCENE_2A</METADATA_PROFILE>
  </Metadata_Id>


  <!-- /// Systeme de coordonnees. /// -->
  <Coordinate_Reference_System>
    <GEO_TABLES version='5.2'>EPSG</GEO_TABLES>
    <Horizontal_CS>
      <HORIZONTAL_CS_CODE>epsg:32622</HORIZONTAL_CS_CODE>
      <HORIZONTAL_CS_TYPE>PROJECTED</HORIZONTAL_CS_TYPE>
      <HORIZONTAL_CS_NAME>WGS 84/UTM22N</HORIZONTAL_CS_NAME>
    </Horizontal_CS>
  </Coordinate_Reference_System>


  <!-- /// Indications d'affichage. /// -->
  <Image_Display>
    <Band_Display_Order>
      <RED_CHANNEL>1</RED_CHANNEL>
      <GREEN_CHANNEL>2</GREEN_CHANNEL>
      <BLUE_CHANNEL>3</BLUE_CHANNEL>
    </Band_Display_Order>
    <!-- ////// -->
  </Image_Display>

</Dimap_Document>
```

Code 2.1 – Un extrait de fichier de métadonnées SPOT 5 en XML. Il contient une information relative au format des données, au niveau de prétraitements, au système de coordonnées, ainsi que des indications d'affichage.

L'image satellitaire fournit l'information de la prise de vue. Les métadonnées de l'image fournissent de nombreuses autres informations qui peuvent être relatives au format ou à l'affichage des données (Code 2.1), concerner la prise de vue ou les paramètres utilisés dans les prétraitements ou bien encore détailler les caractéristiques du satellite et des capteurs (Code 2.2).

Définition 23 (Irradiance ou éclairement énergétique ($\mathrm{W \cdot m^{-2}}$), d'après Pouget (2005))**.** L'*éclairement* d'un objet désigne la puissance du flux électromagnétique qu'il reçoit par unité de surface.

Les caractéristiques de la prise du vue et des instruments embarqués permettent de convertir en mesures de luminance les simples comptes numériques enregistrés (Bonn, 1994).

Définition 24 (Luminance énergétique ($\mathrm{W \cdot m^{-2} \cdot sr^{-1}}$), d'après Pouget (2005))**.** La *luminance* d'un objet désigne la puissance du flux électromagnétique qu'il émet par unité de surface par unité d'angle solide.

[2] http://www.iso.org/iso/catalogue_detail?csnumber=26020 [accessed : 01-02-2013].

```
<?xml version='1.0' encoding='UTF-8'?>
<?xml-stylesheet href="STYLE.XSL" type="text/xsl" ?>

<Dimap_Document xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
 xsi:noNamespaceSchemaLocation='Spot_Scene.xsd' name='METADATA.DIM'>


  <!-- /// Position du satellite /// -->
  <Ephemeris>
    <SATELLITE_ALTITUDE>827720.215571</SATELLITE_ALTITUDE>
    <NADIR_LAT>5.110184</NADIR_LAT>
    <NADIR_LON>-53.023214</NADIR_LON>
    <Points>
      <Point>
        <Location>
          <X>4.4008006345e+06</X>
          <Y>-5.4466700226e+06</Y>
          <Z>1.6938760477e+06</Z>
        </Location>
        <Velocity>
          <X>-2.1200986720e+02</X>
          <Y>-2.4045637900e+03</Y>
          <Z>-7.1402984940e+03</Z>
        </Velocity>
      <TIME>2009-11-02T13:57:56.000000</TIME>
      </Point>
      <!-- ////// -->
    <Points>
    <!-- ////// -->
  </Ephemeris>


  <!-- /// Calibration des capteurs /// -->
  <Sensor_Calibration>
  <!-- ////// -->
    <Band_Solar_Irradiance>
      <BAND_INDEX>1</BAND_INDEX>
      <SOLAR_IRRADIANCE_VALUE>1858</SOLAR_IRRADIANCE_VALUE>
    </Band_Solar_Irradiance>

    <Band_Solar_Irradiance>
      <BAND_INDEX>2</BAND_INDEX>
      <SOLAR_IRRADIANCE_VALUE>1575</SOLAR_IRRADIANCE_VALUE>
    </Band_Solar_Irradiance>

    <Band_Solar_Irradiance>
      <BAND_INDEX>3</BAND_INDEX>
      <SOLAR_IRRADIANCE_VALUE>1047</SOLAR_IRRADIANCE_VALUE>
    </Band_Solar_Irradiance>

    <Band_Solar_Irradiance>
      <BAND_INDEX>4</BAND_INDEX>
      <SOLAR_IRRADIANCE_VALUE>234</SOLAR_IRRADIANCE_VALUE>
    </Band_Solar_Irradiance>
    <!-- ////// -->
  </Sensor_Calibration>
  <!-- ////// -->

</Dimap_Document>
```

CODE 2.2 – Un extrait de fichier de métadonnées SPOT 5 en XML. On y trouve en particulier une information sur la localisation du satellite lors de la prise de vue. Quant à l'information relative à l'éclairement solaire dans chaque bande spectrale, elle est indispensable à la calibration de l'image.

Dans une certaine mesure, le rapport entre la luminance et l'éclairement de la cible permet de s'affranchir des caractéristiques de l'énergie incidente. Il est donc en quelque sorte caractéristique de la cible. Il s'agit de la réflectance.

2.2.4.2 Réflectance

Définition 25 (Réflectance, d'après Pouget (2005))**.** La *réflectance* est définie par le rapport entre la luminance et l'éclairement.

En première approximation, la réflectance est un indicateur relativement caractéristique des surfaces. Il faut néanmoins la manipuler avec précautions.

Remarque 9 (Réflectance spectrique, d'après Bonnet (2005))**.** La notion de réflectance [...] caractérise l'image [...] sans analyse par rapport à la longueur d'onde.

L'étude de la réflectance de nombreux matériaux montre qu'elle dépend de la longueur d'onde. D'où la notion de réflectance spectrique.

Définition 26 (Réflectance spectrique)**.** La *réflectance spectrique* désigne donc le rapport entre la luminance et l'éclairement en fonction de la longueur d'onde.

Remarque 10 (Réflectance spectrique, angle d'incidence, angle d'observation et couleur (Bonnet, 2005))**.** Il est à noter que la fonction de réflectance spectrique est dépendante des angles d'incidence et d'observation, en particulier au voisinage de la réflexion spéculaire. Elle se traduit pour l'œil humain par la notion de couleur. Il en résulte qu'un objet homogène, en particulier tridimentionnel, ne donne pas une couleur homogène pour l'image.

Remarque 11 (Réflectance et réflectivité)**.** D'après la Commission Internationale de l'Éclairage, on distingue la réflectivité de la réflectance par le fait que la réflectivité s'applique à la réflexion en profondeur des objets. Quand la réflexion concerne de fines couches de matériau, les effets de la réflexion interne peuvent provoquer une variation de la réflectance en fonction de l'épaisseur de la surface. La réflectivité est la valeur limite de la réflectance lorsque l'échantillon s'épaissit ; c'est la réflectance intrinsèque de la surface, indépendamment d'autres facteurs tels que la réflectance de la surface arrière. Une autre interprétation est que la réflectance est la fraction de puissance électromagnétique réfléchie par un échantillon particulier, alors que la réflectivité est une propriété du matériau en soi, qui pourrait être mesurée par une machine parfaite si le matériau occupait la moitié de tout l'espace[1].

On le voit, les concepts de la télédétection peuvent être sujets à interprétation et il n'est pas toujours facile de les définir avec précision. Souvent, ils sont employés avec approximation. Ce problème de l'ambiguïté des termes, gérable dans un champ disciplinaire bien délimité, peut devenir plus gênant lorsque différentes disciplines sont amenées à collaborer, chacune avec son propre vocabulaire. C'est pour cette raison que les termes seront définis avec autant de soin que possible.

En ce qui concerne cette section sur la télédétection, retenons simplement en première approximation le sens usuel de la réflectance, indépendamment de la complexité des modèles proposés pour

[1] Traduction réalisée d'après http://en.wikipedia.org/wiki/Reflectivity [accessed : 08-02-2013].

la représenter[1]. C'est le ratio de l'énergie solaire réfléchie par la surface sur l'énergie incidente, aussi appelé albédo[2]. Cette proportion dépend de nombreux paramètres et en particulier de la longueur d'onde.

Sur les images satellitaires, la longueur d'onde et la position du soleil influent sur la quantité d'énergie reçue par la surface et sur la réflectance, non seulement directement, mais également par l'intermédiaire de la taille des ombres, car celles-ci dépendent également de la longueur d'onde et de la position du soleil.

Les réflectances spectriques de nombreuses surfaces naturelles ont été mesurées pour diverses longueurs d'ondes du spectre électromagnétique[3]. Pour un objet observé, la relation étroite quoiqu'imparfaite qui existe entre sa nature et sa réflectance en fonction de la longueur d'onde, conduit au concept de signature spectrale.

2.2.4.3 Signature spectrale

Définition 27 (Signature spectrale)**.** La signature spectrale d'une surface est constituée de valeurs caractéristiques de sa réflectance à différentes longueurs d'ondes. C'est donc un ensemble discret de valeurs, issu de la réflectance spectrique.

Remarque 12. En réalité, les longueurs d'ondes de la signature spectrale sont des *intervalles* de longueurs d'ondes suffisamment étroits pour être la plupart du temps assimilés à une seule longueur d'onde.

Les signatures spectrales permettent de différencier les objets et d'envisager de caractériser leur nature, mais elles varient dans le temps et dans l'espace (Pouget, 2005) en fonction des saisons, des climats... Examinons maintenant les caractéristiques théoriques des surfaces les plus courantes.

Surfaces minérales La réflectance des surfaces minérales dépend de leur composition physico-chimique. On peut distinguer rapidement les deux grandes familles géologique que sont les roches et les sols.

Les roches terrestres sont généralement masquées par les sols. Même lorsqu'elles affleurent, leur patine (altération de surface) et autre lichen modifient leur réflectance intrinsèque.

Les sols sont le produit de l'altération des roches. Leur réflectance diminue lorsqu'ils sont humides et il est important de les reconnaître en télédétection pour pouvoir différencier les sols nus des sols en végétation.

Surfaces végétales Les surfaces végétales sont complexes et changeantes. Chacune a sa propre signature spectrale qui dépend de l'échelle d'observation : feuille, plante, couvert végétal... La réflectance des feuilles dépend de l'espèce, du stade de croissance et de la géométrie de la feuille. La réflectance d'un couvert végétal intègre la réflectance individuelle de la plante et des organes qui la composent. La densité de la couverture végétale et la réflectance du sol interviennent également.

$$\text{indice foliaire} = \frac{\text{surface de feuilles}}{\text{surface du terrain}}$$

En règle générale, l'intégration spatiale et spectrale des signaux en un seul pixel empêche d'isoler la signature des différentes essences. Pour différencier des cultures différentes[4], il faut des images prises à des dates différentes, des vérités terrain et des photointerprètes[5]. On voit ici l'importance de disposer de points de vue complémentaires et de représentations différentes d'un même phénomène. On remarque également le rôle joué dans l'interprétation par l'expérience du photointerprète.

[1] Voir http://domurado.pagesperso-orange.fr/Memoire/ [accessed : 08-02-2013].
[2] Voir http://e-cours.univ-paris1.fr/modules/uved/envcal/html/glossaire.html [accessed : 08-02-2013].
[3] Voir par exemple la bibliothèque spectrale ASTER v2.0 http://speclib.jpl.nasa.gov/ [accessed : 08-02-2013].
[4] Peuplement d'essences monospécifiques.
[5] On appelle *photointerprète* un expert en photointerprétation.

Définition 28 (Vérité terrain)**.** Une *« vérité terrain »* est un constat, une observation, une mesure réalisé *in situ* par un spécialiste du domaine concerné.

La plupart des feuilles vivantes sont caractérisées par une forte absorption dans le visible, la réflectance correspondante est donc globalement faible, avec un maximum dans le vert[1]. Dans le proche infrarouge, la réflectance est au contraire souvent très forte. Elle est toutefois moindre pour les conifères et les végétaux malades et n'est pas caractéristique des végétaux secs ou brûlés. Dans le moyen infrarouge, c'est la teneur en eau du végétal qui joue le plus dans son comportement spectral.

On voit ici que la végétation chlorophyllienne peut être relativement bien décrite en synthétisant les informations d'une forte absorption dans le visible et d'une forte réflectance dans le proche infrarouge, autrement dit par le rapport ou la différence entre le proche infrarouge et une longueur d'onde visible. D'où la notion d'*indice* (Définition 34 page 49).

La présence d'eau L'eau absorbe énormément le rayonnement infrarouge. Les surfaces en eau (ou humides) apparaissent donc très bien dans ce canal, mais il ne donne aucune information sur leur profondeur.

L'eau absorbe moins le rayonnement visible. Ce sont les longueurs d'onde autour de 450 nm qui pénètrent le mieux, et si l'eau est claire, on peut espérer que le signal se réfléchisse sur le fond pour estimer la profondeur. Toutefois, on peut considérer qu'au-delà de 400 m, aucune lumière ne pénètre[2].

La neige (surtout fraîche) a une forte réflectance. Elle est maximale dans le visible et moindre dans l'infrarouge proche et moyen.

L'urbanisation Les espaces urbains sont caractérisés par des surfaces entremêlées de natures différentes : minérales, végétales, métalliques.... Les volumes (bâtiments...) projettent des ombres qui compliquent l'interprétation radiométrique des images.

2.2.4.4 Effets perturbateurs des signatures spectrales

Dérive de la réflectance La réflectance donne une indication intéressante sur la nature les objets, mais comme nous l'avons vu, en tenant compte des conditions de prise de vue, les pixels des images représentent plutôt l'enregistrement d'une luminance.

Sur une même image, on peut considérer que l'éclairement est uniforme et donc que réflectance et luminance sont proportionnelles. Il faut toutefois nuancer cette approximation dans le cas des images de grandes surfaces sur lesquelles la variation de l'éclairement ne peut plus être négligée entre deux points suffisamment éloignés.

Dans tous les cas, entre deux dates différentes, la luminosité varie à cause de multiples facteurs[3], et il faut alors absolument comparer la réflectance des objets sur chacune des images et non pas sa luminance.

Pour avoir une idée de la dérive de la réflectance entre deux images ou deux régions de la même image, il faut se baser sur les différences de luminance de surfaces de référence.

Les effets atmosphériques L'observation de la surface terrestre depuis les satellites implique la traversée de l'atmosphère par les rayonnements incident et réfléchi. Lors de ces passages, des interactions se produisent entre les ondes lumineuses et les composants atmosphériques.

Les effets d'absorption. Les gaz qui composent l'atmosphère absorbent certaines longueurs d'onde.

[1] La notion de *vert* est qualitative et sujette à interprétation. Il s'agit, selon le capteur concerné, de l'image enregistrée dans une longueur d'onde couramment qualifiée de verte.

[2] Voir http://eduscol.education.fr/orbito/pedago/littoral/lh310.htm [accessed : 13-02-2013].

[3] Par exemple, sur les images d'hiver, l'éclairement est plus faible et les ombres plus importantes que sur les images d'été.

La diffusion de Rayleigh. Elle intervient lorsque la taille des particules rencontrées est inférieure à la longueur d'onde. C'est le cas ici des molécules de gaz. Cette diffusion est plus importante dans les faibles longueurs d'onde. C'est la célèbre cause de la couleur bleue du ciel et de l'apparence relativement claire des ombres dans le visible.

La diffusion de Mie. Elle joue son rôle dans le cadre des particules de taille comparable à la longueur d'onde. C'est le cas des aérosols qui concernent l'ensemble des longueurs d'ondes du visible. Les nuages nous apparaissent blancs car les gouttelettes qu'ils contiennent diffusent la quasi-totalité des longueurs d'ondes du visible.

La réfraction. L'onde est déviée lorsqu'elle traverse deux milieux de densité différente.

Les effets du relief Outre l'atmosphère, le relief influe également sur les valeurs enregistrées dans les images (Pouget, 2005) :

- les ombres étant plus sombres dans l'infrarouge que dans le visible, la qualité des canaux infrarouge baisse généralement pour les zones montagneuses (phénomène d'ombre portée) ;
- avec l'altitude, les effets atmosphériques diminuent et les effets dus à la variation de hauteur du soleil sont amplifiés ;
- les effets de pente peuvent faire varier les luminances observées et induire des erreurs de classification dans les images.

2.2.4.5 La calibration radiométrique

Nous avons établi plus haut les notions de réflectance, de signature spectrale, avant d'examiner les principaux effets perturbateurs. Nous allons donc maintenant préciser les mécanismes de correction de la mesure de réflectance caractéristique des objets observés.

Définition 29 (Calibration radiométrique (absolue), d'après Bonn (1994))**.** La calibration radiométrique des données permet de convertir le signal enregistré par le satellite en une variable physique telle que la luminance, exprimée en unités scientifiques fondamentales.

Remarque 13 (Calibration radiométrique)*.* Dans ce document, la calibration radiométrique désignera par défaut un résultat en réflectance.

Remarque 14 (Calibration radiométrique relative, d'après Bonn (1994))*.* [Dans le cas où le capteur d'un canal se compose de plusieurs détecteurs,] l'égalisation des détecteurs d'un même canal a pour objet de réduire le lignage subsistant, par différentes procédures d'égalisation d'histogramme de la scène[1].

La réflectance exoatmosphérique (ou effective au satellite) permet de s'affranchir dans une certaine mesure de la variabilité de l'éclairement solaire, mais la correction des perturbations atmosphériques nécessite des informations météorologiques (Bonn, 1994).

La FIGURE 2.5 représente le processus de calibration de l'image en réflectance. On remarque qu'il se dégage un schéma, représenté en FIGURE 2.6, par lequel une information de base est enrichie par l'intervention d'une information supplémentaire.

À titre d'illustration sur la forme concrète que peut prendre l'information contextuelle, on peut examiner un extrait de métadonnées d'une prise de vue (CODE 2.3). Le gain et le biais sont des caractéristiques liées aux capteurs, la position du Soleil renseigne sur l'éclairement etc.

[1] Pour en savoir plus sur les problèmes de lignage et des propositions de solutions, on pourra consulter Féty and Terré (1991).

```
<?xml version='1.0' encoding='UTF-8'?>
<?xml-stylesheet href="STYLE.XSL" type="text/xsl" ?>

<Dimap_Document xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
 xsi:noNamespaceSchemaLocation='Spot_Scene.xsd' name='METADATA.DIM'>


  <!-- /// Informations sur la prise de vue. /// -->
  <Scene_Source>
    <GRID_REFERENCE>691340</GRID_REFERENCE>
    <SHIFT_VALUE>0</SHIFT_VALUE>
    <IMAGING_DATE>2009-11-02</IMAGING_DATE>
    <IMAGING_TIME>14:00:21</IMAGING_TIME>
    <MISSION>SPOT</MISSION>
    <MISSION_INDEX>5</MISSION_INDEX>
    <INSTRUMENT>HRG</INSTRUMENT>
    <INSTRUMENT_INDEX>2</INSTRUMENT_INDEX>
    <SENSOR_CODE>J</SENSOR_CODE>
    <SCENE_PROCESSING_LEVEL>0</SCENE_PROCESSING_LEVEL>
    <INCIDENCE_ANGLE>6.662696</INCIDENCE_ANGLE>
    <VIEWING_ANGLE>5.873222</VIEWING_ANGLE>
    <SUN_AZIMUTH>137.753870</SUN_AZIMUTH>
    <SUN_ELEVATION>63.467765</SUN_ELEVATION>
    <!-- ////// -->
  </Scene_Source>


  <!-- /// Parametres utilises pour l'interpretation physique. /// -->
  <Image_Interpretation>
    <Spectral_Band_Info>
      <BAND_DESCRIPTION>XS1</BAND_DESCRIPTION>
      <PHYSICAL_UNIT>equivalent radiance (W.m-2.Sr-1.um-1)</PHYSICAL_UNIT>
      <PHYSICAL_BIAS>0.000000</PHYSICAL_BIAS>
      <PHYSICAL_GAIN>1.152462</PHYSICAL_GAIN>
      <PHYSICAL_CALIBRATION_DATE>2009-09-01T00:00:00.000000</
       PHYSICAL_CALIBRATION_DATE>
      <BAND_INDEX>1</BAND_INDEX>
    </Spectral_Band_Info>
    <!-- ////// -->
  </Image_Interpretation>

  <!-- ////// -->
</Dimap_Document>
```

CODE 2.3 – Des informations utiles aux traitements d'image contenues dans des métadonnées SPOT 5.

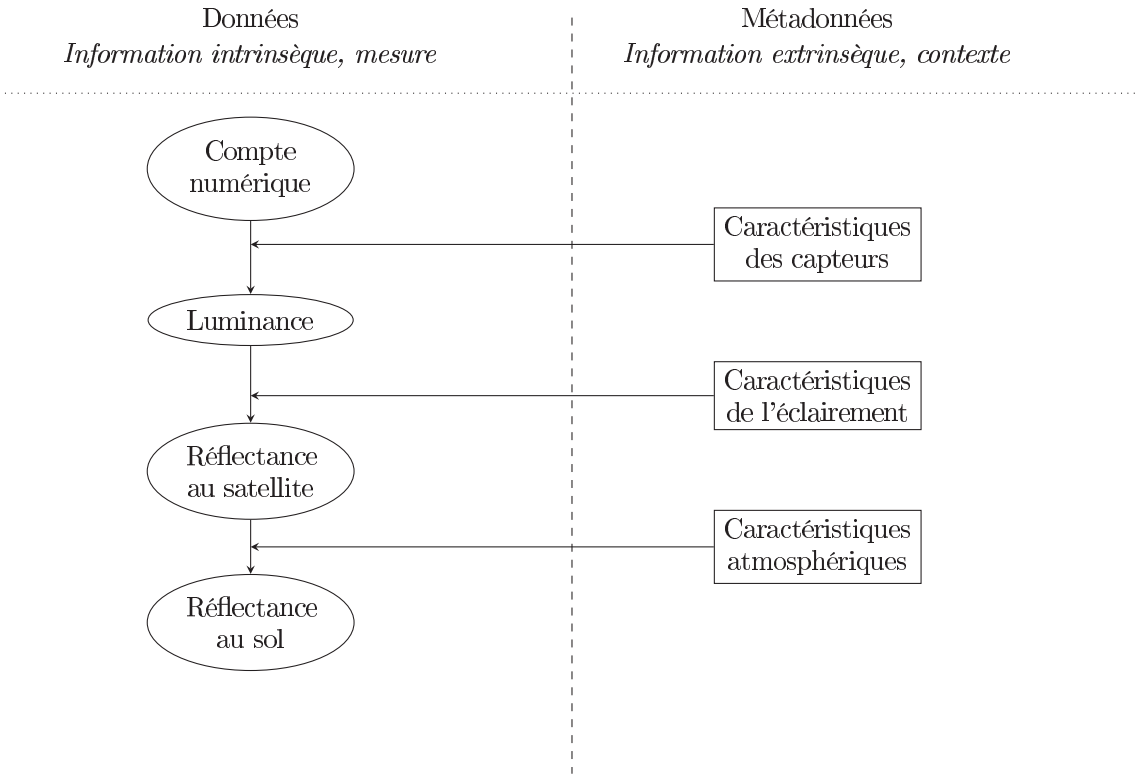


FIGURE 2.5 – Les étapes de calibration radiométrique. L'information principale est représentée dans un ovale, l'information extérieure dans un rectangle. Le CODE 2.3 donne un exemple concret de métadonnées relatives aux caractéristiques des capteurs et de l'éclairement. L'information sur les caractéristiques atmosphériques sont généralement modélisées ou acquises via d'autres instruments. Elle permet d'estimer plus précisément la réflectance « au sol », c'est-à-dire corrigée des perturbations atmosphériques.

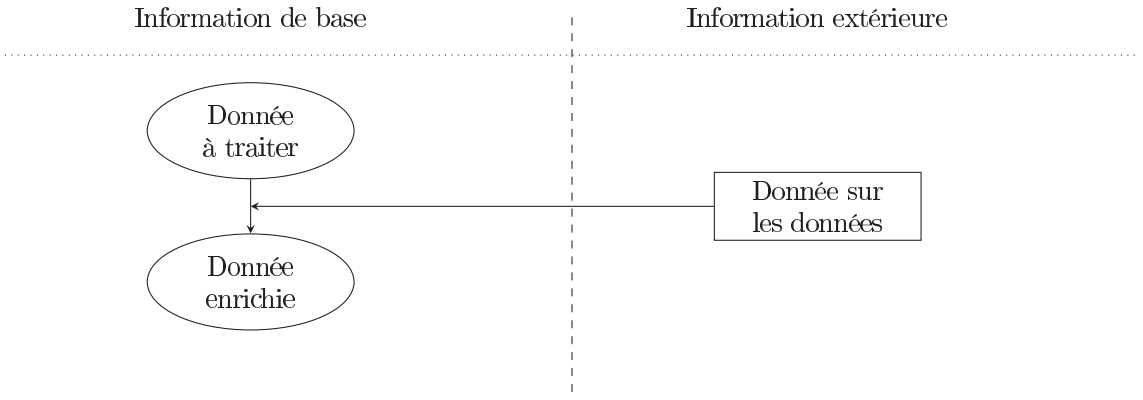


FIGURE 2.6 – Le processus d'enrichissement de l'information. L'information principale est représentée dans un ovale, l'information extérieure dans un rectangle.

2.2.5 Les niveaux de prétraitement des images SPOT

Les corrections géométriques et radiométriques que nous avons évoquées constituent des prétraitements courants en télédétection.

Au cours du développement de nos travaux, nous avons utilisé différentes images : le principe

reste le même. Toutefois, le projet CARTAM-SAT concernait plus particulièrement les images SPOT de Guyane, c'est pourquoi nous allons détailler rapidement, les niveaux de prétraitements proposés par SpotImage[1] :

- le niveau 1A intègre les corrections radiométriques des distorsions dues aux écarts de sensibilité entre les détecteurs élémentaires de l'instrument de prise de vue (calibration radiométrique relative, voir Remarque 14).
- le niveau 1B comprend les corrections radiométriques du niveau 1A auxquelles vient s'ajouter la correction géométrique de certains effets dus aux conditions de prise de vue[2]. Les images du niveau 1B sont adaptées aux calculs de distances, d'angles et d'aires.
- le niveau 2A inclut également les corrections du niveau 1A. Les corrections géométriques sont réalisées dans la projection cartographique standard[3] sans points d'appui.
- le niveau 2B (à la différence du niveau 2A) utilise des points d'appui dans la mise en projection. Les corrections géométriques se basent sur une modélisation de la dynamique de vol du satellite (SPOT-b, 2010).
- le niveau 3 utilise un modèle numérique d'élévation de manière à améliorer les corrections géométriques, particulièrement dans les zones d'altitude élevée.

Bien que cette catégorisation des niveaux de prétraitement soit spécifique aux images SPOT, le principe est transposable à la plus large part des images satellitaires.

2.2.6 Les résolutions d'une image satellitaire

Les notions de résolutions des images numériques matricielles s'appliquent également aux images satellitaires. Toutefois, le vocabulaire de la communauté présente quelques spécificités d'usage dues aux particularités des données.

Définition 30 (Résolution radiométrique)**.** Pour un canal donné, la résolution radiométrique d'une image satellitaire désigne sa résolution tonale[4].

La résolution radiométrique indique le nombre de valeurs que le capteur peut discriminer. Plus la résolution radiométrique est élevée, plus un capteur est capable de différentier de petites variations d'énergies.

Définition 31 (Résolution spectrale)**.** La résolution spectrale d'un capteur décrit sa capacité à utiliser de petites fenêtres de longueurs d'onde. Plus la résolution spectrale est fine, plus les fenêtres des différents canaux du capteur sont étroites.

Des instruments de télédétection appelés *capteurs multispectraux* peuvent enregistrer l'énergie reçue selon des intervalles de longueurs d'onde dans différentes bandes spectrales. Les capteurs encore plus développés, dits *hyperspectraux* sont capables de détecter des centaines de bandes spectrales.

[1] http://www.spotimage.com/web/fr/153-niveaux-de-pretraitement-et-precision-de-localisation.php
[2] La variation d'altitude du satellite, l'effet panoramique, la courbure et la rotation de la Terre.
[3] UTM WGS84 par défaut
[4] Voir la Définition 8.

Remarque 15 (Relation entre les résolutions spatiale, radiométrique et spectrale). Au niveau des capteurs, l'amélioration d'une des trois résolutions spatiale, radiométrique ou spectrale n'est pas sans impact sur les autres et contraint à des compromis en fonction des objectifs de télédétection.

En effet, l'amélioration de la résolution spatiale pousse à réduire la surface des détecteurs d'énergie, donc à réduire l'énergie recueillie. Cela affecte la résolution radiométrique, car avec moins d'énergie la capacité du capteur à détecter des écarts d'intensité énergétique diminue également.

Si l'on veut préserver la quantité d'énergie reçue par le capteur tout en augmentant la résolution spatiale, il faut alors augmenter la fenêtre spectrale. On perd alors en résolution spectrale[1].

2.2.7 La télédétection par les hyperfréquences

Les techniques utilisant les *hyperfréquences* peuvent être actives ou passives. La grande longueur d'onde des micro-ondes par rapport aux domaines visible et infrarouge, leur confère des propriétés particulières.

2.2.7.1 Capteurs hyperfréquences passifs

Ces instruments sont utilisés en météorologie (quantités d'eau et d'ozone dans l'atmosphère), hydrologie, océanographie (cartographie, glaces de mer, courants et vents de surface ou encore détection de polluants (nappes d'hydrocarbures).

L'énergie de ces longueurs d'ondes est faible et le champ de vision doit donc être suffisamment grand pour récupérer l'énergie nécessaire, ce qui dégrade en général la résolution spatiale.

2.2.7.2 Capteurs hyperfréquences actifs

Le plus connu est le RADAR (*Radio Detection And Randging*, soit : détection et télémétrie par ondes radio). Un RADAR émet vers la cible un signal dans les hyperfréquences et détecte sa partie rétrodiffusée. L'intensité du signal reçu renseigne sur la nature de la cible et le délai entre l'émission et la réception renseigne sur sa portée (distance).

Ces capteurs peuvent être classés en deux catégories :

- Les capteurs non-imageurs, tels les altimètres et les diffusomètres, ne produisent pas d'image.
- Les capteurs imageurs produisent des images dans presque toutes les conditions atmosphériques, car les ondes ne subissent pas la diffusion provoquée par les nuages, la bruine, la poussière et la pluie fine.

2.3 Enregistrement, traitement et analyse de l'information des images de télédétection

En dépit de la spécificité des images satellitaires, les traitements qui s'y appliquent ne sont la plupart du temps que l'adaptation ou l'extension de traitements sur les images en général.

2.3.1 Nature de l'image de télédétection

Nous avons vu plus haut que l'image satellitaire est une image numérique matricielle, mais que contrairement aux photographies, l'ensemble des valeurs qui la composent peuvent présenter les unes par rapport aux autres un certain décalage dans le temps. Cela dépend du capteur utilisé.

[1] Remarque issue du site de *Ressources Naturelles Cananda* http://www.rncan.gc.ca/sciences-terre/limite-geographique/teledetection/fondements/1775 [accessed : 14-02-2013].

Une partie importante de l'information en télédétection est déduite de la longueur d'onde. Des images d'un même lieu à des longueurs d'ondes différentes permettent certaines déductions sur la nature des surfaces représentées, grâce à la connaissance de signatures spectrales. Pour cela, de nombreux satellites d'observation de la Terre embarquent des capteurs imageurs multispectraux, voire hyperspectraux, permettant l'acquisition d'images multispectrales. Nous avons également vu que les images de télédétection sont potentiellement caractérisées par trois résolutions différentes entre lesquelles il est nécessaire de trouver un compromis.

Les valeurs d'une image multispectrale peuvent être représentées de différentes manières selon l'information que l'on souhaite faire ressortir (Landgrebe, 1997). Toutes ces caractéristiques font de l'imagerie satellitaire une discipline particulière à interpréter, traiter et visualiser en connaissance de cause.

2.3.2 Format des images

Les techniques d'enregistrements des données sont variées[1]. Les formats de base BIL, BIP et BSQ permettent d'optimiser certaines méthodes d'accès aux données plutôt que d'autres, suivant si l'on souhaite obtenir, pour un pixel, l'ensemble de ses valeurs dans différentes bandes, ou bien les valeurs de tous les pixels dans une seule bande[2].

On peut également remarquer l'existence de méthodes sophistiquées destinées à optimiser l'espace de stockage. RLE (*Run-Length Encoding*)[3] en fait partie, qui présente un intérêt pour les images homogènes, car les valeurs identiques de pixels contigus sur une ligne ne sont enregistrées qu'une fois, avec l'indication du nombre de répétitions.

En général, un format graphique contient, d'une part les métadonnées sur l'image dans un en-tête, et d'autre part les données de l'image proprement dite.

De fait, les images satellitaires peuvent être enregistrées sous de nombreux formats courants d'images matricielles comme JPEG 2000 (norme ISO 15444-1) ou TIFF. Une déclinaison particulière de ce dernier format, GeoTIFF, permet d'intégrer directement au fichier de l'image des métadonnées utiles au géoréférencement.

Les images proposées par *Spot Image*[4] sont au format Dimap, structuré de la manière suivante[5] :

- d'une part l'image proprement dite, au format *GeoTIFF*, composé :
  - des données de l'image au sens strict, enregistrées au format TIFF ;
  - des informations de géoréférencement contenues dans la partie géographique de l'en-tête du fichier.
- d'autre part les données de l'image relatives à l'acquisition et aux prétraitements, enregistrées dans un fichier utilisant une syntaxe XML.

Le format Dimap permet à Spot Image de réduire le nombre de fichiers associés à l'image tout en se préparant à adopter la norme de description de métadonnées ISO 19115 (SPOT-a, 2010).

2.3.3 Mise en évidence visuelle par rehaussement et filtrage, traits des images

L'information contenue dans une image est très volumineuse et il peut être pratique de la résumer de manière à en acquérir rapidement un aperçu global ou bien à faire ressortir des aspects particuliers.

[1] On trouvera une description des principaux formats à l'adresse http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?topicname=BIL,_BIP,_and_BSQ_raster_files [accessed : 15-02-2013].

[2] On se repportera utilement à la page http://geol.hu/data/online_help/ENVI_File_Formats.html [accessed : 15-02-2013].

[3] Aussi appelé RLC (*Run-Length Coding*).

[4] *Spot Image* assure l'acquisition et la diffusion des données SPOT.

[5] http://www.spotimage.com/web/fr/154-le-format-dimap.php. [accessed : 06-05-2013]

2.3.3.1 Histogramme et rehaussement

Une manière de synthétiser l'information de l'image est d'en calculer l'histogramme.

Définition 32 (Histogramme)**.** L'histogramme d'une image est la représentation graphique du nombre des pixels d'une image en fonction de leur valeur. La fonction discrète h de l'histogramme est définie par :

$$h(n) \triangleq \sum_{x=1}^{largeur} \sum_{y=1}^{hauteur} \chi_{\{n\}}(p(x,y))$$

où n est une valeur possible pour les pixels et $\chi_{\{n\}}$ est la fonction caractéristique sur l'ensemble des pixels de l'image. $p(x,y)$ est la valeur du pixel de coordonnées (x,y) dans l'image.

On s'aperçoit alors que les valeurs de l'image n'exploitent pas forcément toute l'étendue de la plage dynamique, ce qui se traduit visuellement par un contraste améliorable.

La méthode la plus simple pour faciliter l'interprétation visuelle des images et leur compréhension est le rehaussement linéaire. Il s'agit tout d'abord d'identifier les bornes inférieure et supérieure des valeurs des pixels puis d'étirer les valeurs linéairement de manière à utiliser l'ensemble des bits disponibles pour leur codage.

Il existe toutes sortes d'autres méthodes non-linéaires mais dont la finalité est identique.

(A) Fausses couleurs (B) Bande « verte » $(0,5-0,59\ \mu m)$ (C) Bande « verte » (après rehaussement)

FIGURE 2.7 – Exemple de rehaussement d'image. En (B), un canal représenté en niveaux de gris est extrait de l'image multispectrale d'origine (A) représentée ici en fausses couleurs. Mais les données du canal en question n'exploitent pas toutes les valeur possibles (voir FIGURE 2.8). En recalculant les valeurs des pixels de manière à exploiter l'ensemble des valeurs de 0 à 255 (voir FIGURE 2.9), l'image (C) montre que l'on améliore le contraste visuel.

2.3.3.2 Corrélogramme de couleurs

Une autre technique, plus riche mais également plus complexe que l'histogramme, permet de représenter la corrélation spatiale des couleurs dans une image $\mathfrak{I}$. Pour une couleur c_1 donnée, le *corrélogramme* de couleurs donne la probabilité qu'un pixel situé à la distance d soit de couleur c_2.

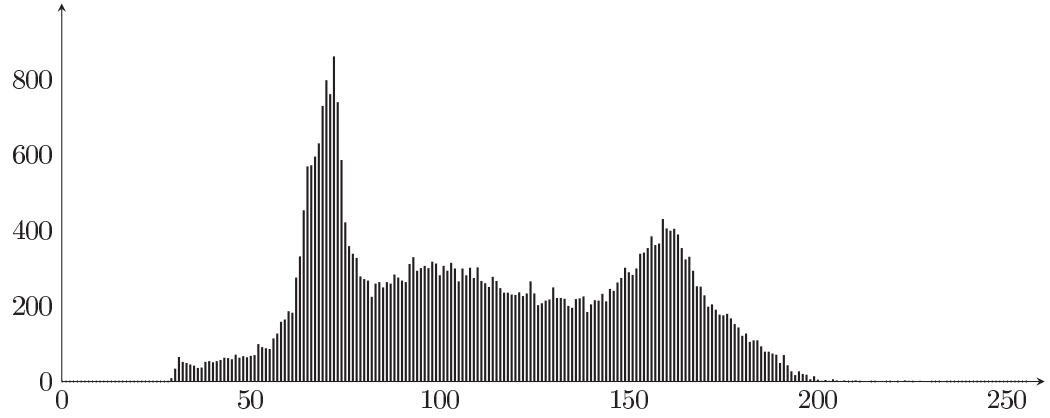


FIGURE 2.8 – Exemple d'histogramme : la répartition des pixels en fonction de leur valeur (de 0 à 255) dans l'image (B) de la FIGURE 2.7. À l'observation des extrémités de l'histogramme, on constate qu'aucun pixel n'a de valeur proche de 0 ou de 255. À l'affichage, cela se traduit par l'absence de blanc et de noir.

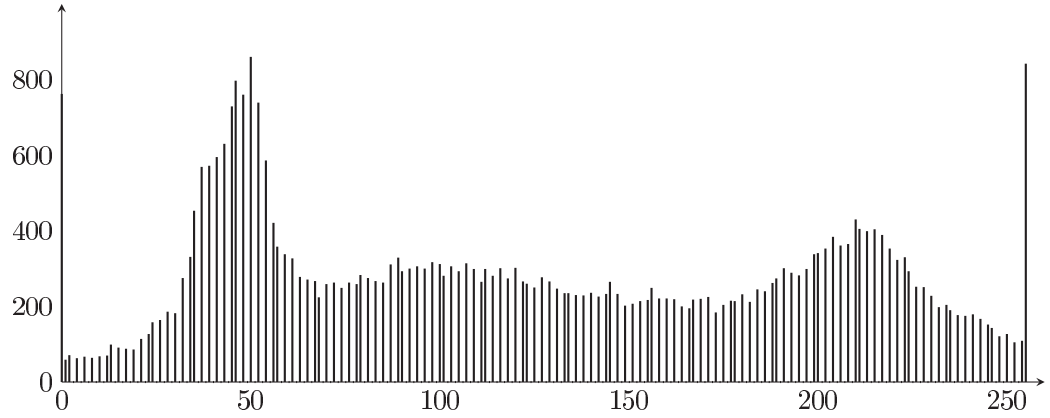


FIGURE 2.9 – Exemple d'histogramme : l'effet du rehaussement linéaire sur l'histogramme dans l'image (C) de la FIGURE 2.7 (à comparer avec l'histogramme en FIGURE 2.8).

Définition 33 (Corrélogramme, d'après Xiaoguang (2002))**.** Avec P une loi de probabilité, p et q deux pixels séparés dans l'image par la distance d, et $\mathfrak{I}_{n_i}$ l'ensemble des pixels de l'image de couleur n_i, on a la définition suivante :

$$\gamma_{\mathfrak{I}}(d, n_1, n_2) \triangleq \mathrm{P}[p \in \mathfrak{I}_{n_2} | (||p - q|| = d \wedge q \in \mathfrak{I}_{n_1})]$$

C'est-à-dire que le corrélogramme de l'image $\mathfrak{I}$ est défini par la probabilité qu'un pixel soit de couleur n_2, sachant qu'il se trouve à la distance d d'un pixel de couleur n_1.

2.3.3.3 Le filtrage

Les filtres spatiaux sont une méthode de traitement numérique. Ils sont conçus de manière à faire ressortir des caractéristiques spécifiques comme par exemple les éléments directionnels. Le principe est le suivant :

- une fenêtre de quelques pixels de côté se déplace sur l'image ;

- un traitement mathématique utilise les pixels sous la fenêtre pour affecter une nouvelle valeur au pixel central.

On peut citer quelques exemples classiques de ce principe. Les filtres passe-bas réduisent les petits détails d'une image et accentuent les zones homogènes (filtres moyenneur et médian utilisés avec les images radar). Au contraire, un filtre passe-haut fait ressortir les petits détails.

Les filtres directionnels ou filtres de détection de contours sont utilisés pour rehausser les caractéristiques linéaires d'une image (routes, limites de parcelles agricoles, structures géologiques...)

2.3.4 Opérations sur les images

Les opérations sur les images génèrent une nouvelle image à partir de différentes sources d'information (Pouget, 2005).

2.3.4.1 Soustraction et division d'images

La soustraction d'images est souvent utilisée dans la détection de changements. Il s'agit de calculer la valeur des pixels p_3 de coordonnées (x, y) dans l'image résultat, à partir des pixels p_1 et p_2 de mêmes coordonnées.

$$p_3(x, y) = p_1(x, y) - p_2(x, y) + C$$

On peut recentrer l'image résultat en ajoutant une constante à l'ensemble des pixels, si l'on veut par exemple éviter les valeurs négatives.

La division d'images permet de mettre en évidence des surfaces d'occupation du sol en utilisant leurs différences de réponse dans deux bandes spectrales différentes.

2.3.4.2 Indices spectraux

La différence et la division d'images sont des opérations très courantes en télédétection. Elles servent en particulier à définir les indices spectraux, caractéristiques d'une occupation du sol particulière, comme la végétation chlorophyllienne (FIGURE 2.10). En général la végétation réfléchit fortement le proche infrarouge et absorbe fortement le rouge visible.

Définition 34 (Indice spectral)**.** Un indice spectral est une combinaison d'opérations (addition, soustraction, division et multiplication) sur des images spectrales permettant de mettre en évidence une information particulière.

Définition 35 (Image pseudospectrale)**.** On appellera image pseudospectrale l'image résultant du calcul d'un indice spectral.

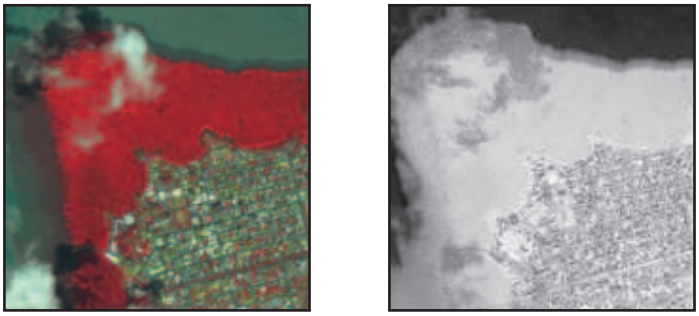
Quelques indices spectraux de référence sont indiqués en Annexe A.1.

2.3.5 Notion de texture

2.3.5.1 Origines de la notion de texture

La notion de texture est assez intuitive et difficile à définir avec rigueur. Elle ne présuppose pas forcément d'arrangement répétitif d'une primitive structurelle et peut également être organisée de manière aléatoire. La texture est un élément très important dans la reconnaissance visuelle (Bissol, 2006).

Plusieurs techniques sont recensées pour l'exploitation des textures de l'image. Une première approche *statistique* se base principalement sur des matrices de co-occurrence des valeurs des pixels (Bissol, 2006). Une autre, dite *géométrique*, repose sur la recherche de composants élémentaires de



(A) Fausses couleurs (B) Indice de végétation (NDVI)

FIGURE 2.10 – Exemple de calcul d'indice (B) sur une image multispectrale de Cayenne (A). L'indice en question est le NDVI (*Normalized Difference Vegetation Index*) calculé à partir des bandes rouge et proche infrarouge (c'est-à-dire les canaux 2 et 3 précisés en TABLE 2.2). La couverture végétale (en rouge sur l'image en fausses couleurs) correspond à un fort NDVI : les zones les plus claires sur l'image (B). Les zones les plus sombres qui correspondent à de l'eau ont un NDVI très faible. Dans la zone hétérogène (urbanisée) le NDVI n'est pas négligeable. Dans l'absolu, cela pourrait correspondre à la présence de végétation chlorophyllienne en ville ou bien simplement à un pouvoir séparateur moindre du NDVI entre la végétation et l'urbanisme qu'entre la végétation et l'eau.

la texture, ce qui semble assez peu adapté aux textures naturelles. Enfin, les approches *fréquentielles* exploitent les fréquences de l'image, sachant que les surfaces « lisses » reflètent des variations plus graduelles que les surfaces d'aspect plus « rugueux ».

D'après Bissol (2006), aucune méthode de calcul de texture ne fait consensus. Toutefois la bibliothèque *Orfeo Toolbox*[1], comme d'autres programmes, implémente des algorithmes basés sur la matrice de co-occurrence. Nous allons donc brièvement nous attarder sur la signification de cette technique.

2.3.5.2 Matrice de co-occurrence en niveaux de gris[2]

Définition 36 (Matrice de co-occurrence en niveaux de gris, d'après Haralick et al. (1973))**.** La matrice de co-occurrence en niveaux de gris (*Grey Levels Co-occurrence Matrix* – GLCM) est applicable à une image en niveaux de gris pour laquelle elle représente la fréquence des combinaisons de niveaux de gris dans l'image[3].

Couramment, on considère des matrices de co-occurrence du *second ordre*, c'est-à-dire représentant des combinaisons de deux pixels, typiquement un pixel et son voisin dans une *direction d'adjacence* donnée[4].

Considérons une image matricielle de résolution tonale égale à 4 dont les valeurs des pixels sont représentées dans le tableau suivant :

$$\mathfrak{I} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

[1]Il s'agit de la bibliothèque logicielle utilisée dans la mise en œuvre pour le traitement des images satellitaires.

[2]La suite de la présentation de la matrice de co-occurrence et des textures qui en sont issues sont principalement des simplifications du tutoriel situé à l'adresse http://www.fp.ucalgary.ca/mhallbey/tutorial.htm, lui-même basé sur la les exemples proposés par Haralick et al. (1973).

[3]D'après le tutoriel http://www.fp.ucalgary.ca/mhallbey/tutorial.htm. Dans la publication d'Haralick et al. (1973), cette matrice est appelée « Matrice de dépendence spatiale en niveaux de gris » (*Gray-Tone Spatial-Dependence Matrix*).

[4]Cette direction d'adjacence revêt une grande importance dans le résultat final de la matrice de co-occurrence.

On peut décrire les co-occurrence de niveaux de gris de cette image dans une direction et un sens donnés. Les valeurs du tableau suivant indiquent, dans la case de coordonnées (i, j), le nombre de pixels de valeur j situé immédiatement à droite d'un pixel de valeur i.

$$\mathfrak{O}_{droite} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

En transposant cette matrice, on obtient les occurrences à gauche. L'addition de ces deux matrices produit une matrice diagonale de co-occurences horizontales nécessaire aux calculs de texture. En normalisant le résultat, on obtient la GLCM *horizontale*. Par un processus similaire, on peut établir une matrice *verticale* et des matrices *diagonales*.

$$\mathfrak{H} = \frac{1}{12} \cdot \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Remarque 16 (Résolution tonale et taille de la GLCM). La taille de la fenêtre de calcul d'une GLCM n'influe pas sur sa taille fixée au carré du nombre de valeurs possibles pour chaque pixel. 256 valeurs possibles produisent donc des GLCM de grande taille, souvent creuses. Il est donc courant de réduire l'éventail des valeurs possibles de l'image préalablement au calcul de la matrice de co-occurrence de manière à améliorer les performances tout en gardant une information encore significative[1].

La matrice GLCM est à la base du calcul des 14 *textures d'Haralick*. Dans l'exemple ci-dessus, les indices de texture calculés à partir de cette matrice sont une information globale de l'image. Mais il est également possible de calculer la GLCM sur de petites régions de l'image de manière à rendre compte de l'hétérogénéité locale dans l'image. Lorsque la fenêtre de la GLCM s'apparente à un voisinage de chaque pixel, on peut alors produire des images de textures (par exemple, l'*inertie*[2] en FIGURE 2.11).

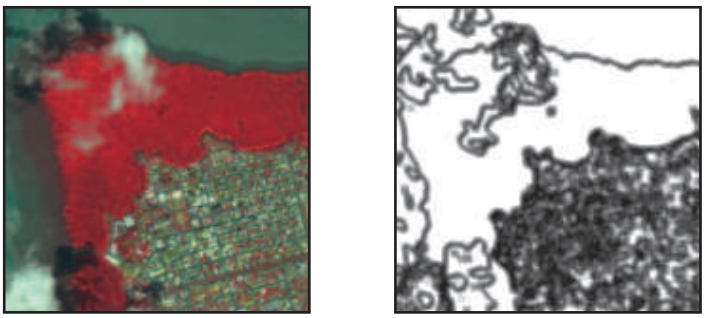
Définition 37 (Image texturale)**.** On appellera image texturale l'image résultant du calcul d'une texture au voisinage de chaque pixel de l'image d'origine.

Quelques indices de texture de référence sont indiqués en Annexe A.2.

Remarque 17. Une texture basée sur la GLCM dépend :

- de *taille de la fenêtre* de calcul de la GLCM ;
- du *voisinage* dans lequel les co-occurences sont comptabilisées autour de chaque pixel ;
- de la *direction* dans laquelle les co-occurences sont comptabilisées ;
- de la *formule* de calcul de texture à partir de la GLCM.

[1] D'après le tutoriel http://www.fp.ucalgary.ca/mhallbey/tutorial.htm.
[2] L'inertie désigne ici l'une des formules de textures d'Haralick.



(A) Fausses couleurs (B) Inertie (du moyen infrarouge)

FIGURE 2.11 – Exemple de calcul de texture (B) sur une image multispectrale de Cayenne (A). La texture en question est une *inertie* calculée sur le moyen infrarouge : les zones foncées (faible inertie) sur l'image (B) peuvent être mises en correspondance des zones hétérogènes de l'image d'origine (A). Inversement, l'inertie des zones homogènes est plus élevée.

2.3.6 Techniques de classification appliquées aux images

Définition 38 (Classification)**.** Classifier un élément signifie lui *attribuer une classe*. Dans le cadre des images satellitaires, la classification peut s'appliquer à trois objets différents. En imagerie, il est possible de classifier les pixels individuellement, des groupes de pixels constitués ou bien encore l'image dans son ensemble.

Généralement, la classe d'appartenance de chaque pixel est définie en fonction de sa valeur dans une ou plusieurs bandes spectrales (Pouget, 2005). La classe peut prendre une forme thématique (eau, maïs...) et la classification constitue donc une première approche méthodologique de rapprochement entre l'image et la sémantique.

Comme toutes les méthodes de classification, celles appliquées aux images peuvent être supervisées ou non supervisées.

2.3.6.1 Classification supervisée

En classification supervisée, les classes sont prédéterminées explicitement. Appliquées aux pixels, ces méthodes suivent traditionnellement le protocole suivant :

1. L'analyste commence par identifier des échantillons homogènes d'une image représentatifs de certaines classes. Ces échantillons forment un ensemble de données-tests.
2. Un algorithme détermine la « signature numérique » de chaque classe, c'est-à-dire détermine les caractéristiques les plus discriminantes des surfaces échantillonnées.
3. Enfin, chaque pixel est affecté à la classe qui *se rapproche* le plus de ses propres caractéristiques.

Des méthodes similaires sont *a priori* applicables à toutes sortes d'éléments et en particulier à des groupes de pixels constitués et à des images entières, à partir du moment où l'on peut en extraire des caractéristiques sur lesquelles appliquer la classification.

2.3.6.2 Classification non supervisée

Dans ce genre de classification, c'est un algorithme qui regroupe les pixels de l'image en classes sans que l'analyste ne spécifie les caractéristiques des classes, mais simplement des critères tels que leur nombre, la distance interclasses, la variance intraclasse...

Ces méthodes donnent souvent l'impression d'une plus grande objectivité car elles réduisent l'influence de l'expert dans la définition des classes.

Ce type de classification « en aveugle », moins contraint par l'expertise humaine, a ses propres caractéristiques. Il peut produire des classes distinctes là où l'analyste en voit une seule, ou bien au contraire regrouper en une seule classe des éléments que l'analyste n'aurait pas rapprochés. Cela peut être vu comme des inconvénients ou bien venir éclairer des informations qui seraient passées inaperçues.

2.3.7 Segmentation

Définition 39 (Segmentation, d'après Castilla and Hay (2008))**.** Une *segmentation* est une partition de l'ensemble de l'image en un ensemble de régions mutuellement disjointes.

Remarque 18 (Segmentation)**.** D'après Castilla and Hay (2008), les régions issues d'une segmentation présentent une certaine uniformité interne qui les différencie des régions adjacentes. Même s'il est très courant, ce critère se discute et on ne le retiendra pas ici, considérant qu'une segmentation peut-être absolument arbitraire.

Définition 40 (Image de segmentation)**.** On appellera *image de segmentation* le résultat de l'opération de segmentation sous forme d'image matricielle dans laquelle la valeur associée à chaque pixel est l'identifiant de sa région d'affectation.

Remarque 19 (Image de segmentation et vectorisation)**.** Une image de segmentation, quoique matricielle, est très proche d'une image vectorielle du point de vue de l'information contenue, puisque celle-ci se résume à un contour et un identifiant pour chaque région.

Définition 41 (Segment)**.** Un *segment* est un groupe de pixels issu d'une segmentation et doté d'un identifiant dans l'image.

Définition 42 (Objet de l'image, d'après Castilla and Hay (2008))**.** Un *objet de l'image* est un segment produit par un bon algorithme de segmentation (une *bonne segmentation*).

Remarque 20 (Bonne segmentation)**.** Nous avons vu que selon Castilla and Hay (2008), la segmentation a pour objectif de découper l'image en régions cohérentes pour un expert. Une *bonne segmentation* est donc une segmentation qui parvient plus ou moins à cet objectif, mais cette qualité doit être établie par un expert ou un consensus d'experts.

De même, nous considérerons ici qu'une bonne segmentation produit des régions cohérentes (objets de l'image) et candidates à recevoir une sémantique de télédétection.

Par exemple, l'identification de régions dans l'image dont les contours seraient définis manuellement par un expert en télédétection en fonction de ses objectifs précis, serait une « bonne segmentation ».

La qualification de « bonne segmentation » est subjective et permet de conserver une certaine souplesse dans l'évaluation d'une segmentation. On l'utilisera en particulier pour signifier que l'on considère qu'elle a produit des segments porteurs de sens.

Définition 43 (Entité de l'image)**.** Une *entité de l'image* est un segment d'image ou un pixel.

Définition 44 (Entité (ou objet) de terrain (ou géographique))**.** Une *entité de terrain* est un phénomène localisé du monde qui ne doit pas être confondu avec sa représentation dans l'image (objet de l'image).

Définition 45 (Entité géoréférencée)**.** Une *entité géoréférencée* est une entité géographique ou une entité de l'image reliée à des coordonnées cartographiques ou géographiques référençant sa localisation (si c'est une entité géographique) ou la localisation du phénomène qu'elle représente (si c'est une entité de l'image).

Remarque 21 (Segmentation et classification des pixels)*.* On peut rapprocher les opérations de segmentation et de classification par pixels, car toutes deux produisent finalement des groupes de pixels. Toutefois, il y a trois grandes différences.

Une différence structurelle tout d'abord, puisque la segmentation produit des groupes de pixels connexes, alors que cela n'est pas forcément le cas d'une classification. Une différence téléologique ensuite, puisque la finalité de la classification est en principe d'attribuer une sémantique à l'étiquette des classes, alors que la segmentation attribue des étiquettes d'identification. Une différence ontologique (dans le sens métaphysique du terme) enfin, puisque la segmentation donne naissance à de *nouvelles entités* (les segments) dont l'existence propre est attestée par leur identifiant. Une classification se contente de regrouper des entités existantes sans en créer de nouvelles.

Ces différences n'empêchent pas d'obtenir une segmentation à partir d'une classification.

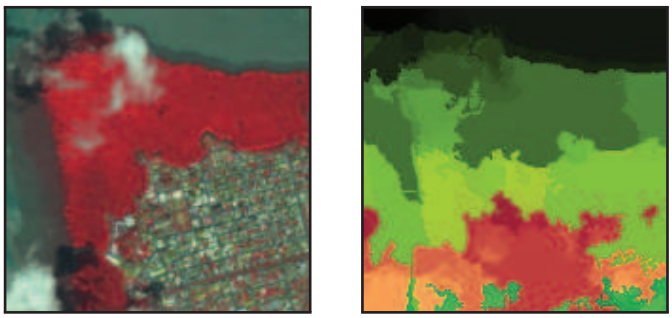
Les méthodes de segmentation sont nombreuses et peuvent différer dans leur approche de la question.

2.3.7.1 La segmentation par régions

Les principaux procédés cités de cette famille sont la *croissance par régions* et la *division-fusion*. Globalement, ces techniques consistent à regrouper des pixels adjacents suivant un *critère d'homogénéité* de manière à former les régions de l'image. Toutefois, la définition des critères de fusion-division n'est pas réputée aisée et le résultat final montre souvent des régions peu précises qui ne correspondent pas forcément aux objets (Sebari and He, 2007).

2.3.7.2 La détection de contours

Une autre manière de voir les choses consiste à détecter les contours des entités figurant sur l'image. On utilise pour cela des opérateurs plus où moins complexes destinés à repérer des changements brutaux dans l'image. La technique de détection des contours présente l'inconvénient de ne pas systématiquement fermer les contours ou bien de segmenter à mauvais escient (Sebari and He, 2007).



(A) Fausses couleurs (B) Segmentation arbitraire

FIGURE 2.12 – Exemple de segmentation (arbitraire) d'une image multispectrale de Cayenne (A). La valeur d'un pixel de l'image en sortie de la segmentation correspond à l'identifiant du segment d'image dont il fait partie. Le résultat peut donc être rigoureusement visualisé en niveaux de gris. La coloration des segments (présentée ici) permet toutefois d'améliorer la visualisation. Les premières difficultés de l'étape de segmentation apparaissent : alors que la composante urbaine de l'image apparaît par exemple bien plus hétérogène que les régions végétalisées, les segments découpés dans ces deux espaces ne présentent pas de différence de taille caractéristique des principaux environnements qui apparaissent sur l'image : eau, végétation, urbanisme...

2.3.7.3 Algorithme de partage des eaux (*watershed*) : une approche mixte

Il s'agit d'un algorithme particulier qui se base généralement sur un gradient de l'image. Les minima locaux sont considérés comme les points de départ des régions. On peut illustrer le fonctionnement de cet algorithme en assimilant l'image du gradient à un relief vallonné[1]. Tout se passe ensuite comme si de l'eau remplissait ce relief. Les régions naissent comme des bassins dans les points les plus bas et fusionnent lorsque le niveau atteint un point-col. La fusion des régions est contrôlée par le paramètre de profondeur, défini comme la différence entre la valeur du point minimum de la surface et le moins élevé des points de sa frontière. Les régions de trop faible profondeur sont agrégées à une région voisine plus profonde, comme si elles débordaient (Team, 2010). Cet algorithme populaire présente une certaine parenté avec les deux approches précédentes.

2.3.7.4 Algorithme de déplacement de la moyenne (*mean shift*) : partition statistique

Meanshift est un autre algorithme utilisé pour la segmentation. Initialement destiné à la classification de points, il regroupe les pixels en fonction de la répartition statistique de leurs valeurs, autour des régions de densité maximale. La conséquence immédiate est que la segmentation d'une zone de l'image dépend également autres zones, dans l'ensemble de l'image considérée (FIGURE 2.13).

2.3.7.5 Autres algorithmes

On retrouve d'autres algorithmes répandus, en particulier l'algorithme de *segmentation multi-résolution* proposé par le logiciel *eCognition*. Cette approche qui produit des résultats souvent intéressants illustre toutefois le problème de la segmentation.

La plupart du temps, les algorithmes ne sont pas compris des utilisateurs, soit parce qu'ils sont complexes, soit parce que les paramètres sont abscons, soit parce que les algorithmes sont propriétaires et fermés. Il en résulte que la segmentation n'est pas guidée par une connaissance explicite de l'utilisateur, mais par une certaine intuition due à l'habitude du programme et complétée par une démarche par essai/erreur[2].

Enfin, les algorithmes sont en quelque sorte prisonniers des paramètres. Par exemple, il n'est pas toujours facile de les ajuster de manière à délimiter simultanément des objets de grande et

[1] Les minima représentent les vallées et les maxima les collines.

[2] On peut consulter le travail de Bertrand (2013) sur l'optimisation de l'unique paramètre d'un algorithme de segmentation en fonction des classes d'occupation du sol recherchées.

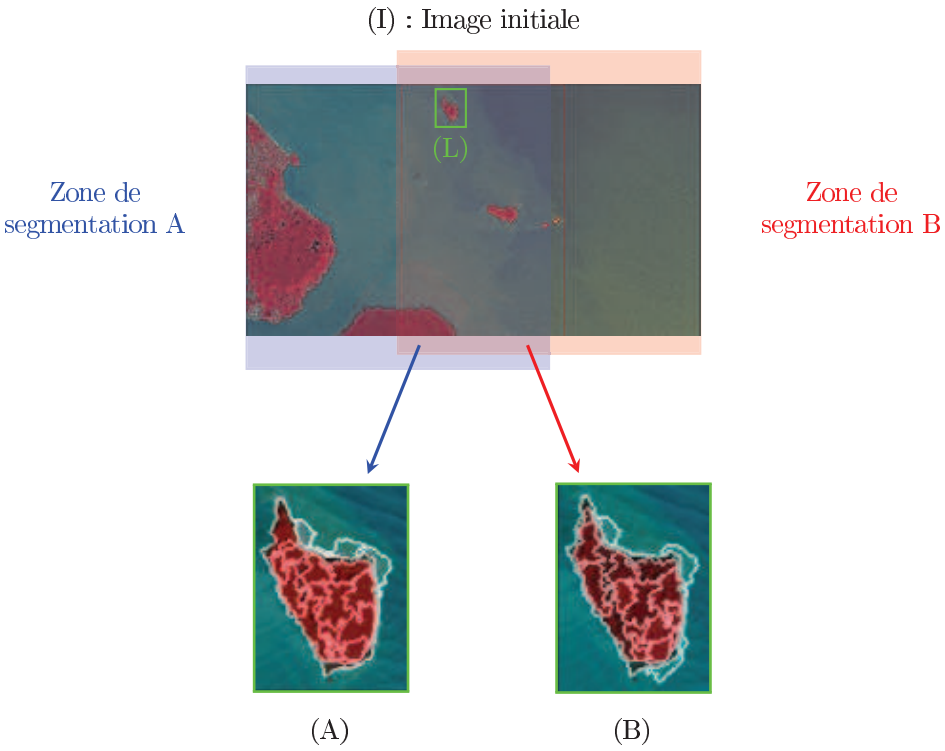


FIGURE 2.13 – Impact du contexte général de l'image sur la segmentation au niveau local avec l'algorithme de déplacement de la moyenne. L'algorithme est appliqué deux fois, avec des paramètres identiques, mais sur deux extraits différents de l'image (I) partageant une étendue commune. Une observation locale de la zone (L) située dans l'étendue commune, montre que le découpage de l'image diffère, pour une même zone (avec un algorithme et des paramètres identiques) en fonction de l'extrait d'image considéré. Cet exemple illustre la difficulté que peut poser la segmentation dans l'association des objets de l'image avec des caractéristiques qui leur seraient propres. En d'autres termes : *comment définir correctement un objet de l'image dont les contours dépendent de pixels représentant des régions situées à des dizaines de kilomètres de distance ?*

de petite taille. Cette remarque vaut également pour les algorithmes de partage des eaux et de déplacement de la moyenne (FIGURE 2.14).

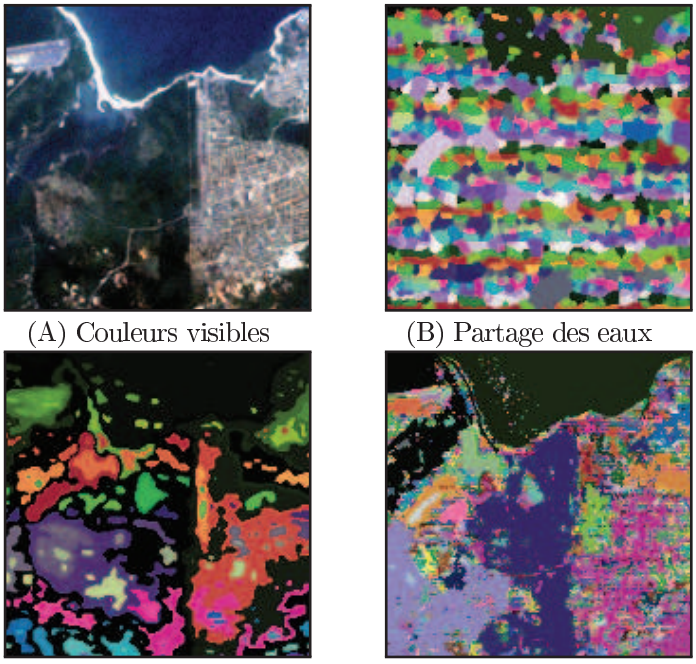
2.3.7.6 Segmentation par seuillage

On peut également segmenter l'image en se basant sur un seuillage des pixels, soit directement appliqué à l'image d'origine, soit après un traitement des pixels.

En particulier, il est possible d'utiliser le seuillage sur une préclassification des pixels. Cette technique présente deux avantages. Premièrement, le mécanisme de partitionnement de l'image est transparent (une préclasse peut être reliée à une sémantique de télédétection) et tranche avec la difficulté d'interprétation des paramètres des autres algorithmes. Deuxièmement, le seuillage produit des segments uniquement à partir de leurs caractéristiques intrinsèques (les pixels qui les composent), par opposition à l'algorithme *meanshift* par exemple, dans lequel l'ensemble de l'image influe sur la définition de chaque segment.

Intuitivement, ce dernier point est assez satisfaisant si l'on considère que la finalité d'un segment est d'être considéré comme un objet de l'image, c'est-à-dire, l'*image d'un objet*. Dès lors, on ne voit pas pourquoi le découpage d'un segment dans l'image devrait être influencé par des pixels à l'autre bout de l'image.

Cela est d'autant plus vrai en télédétection lorsqu'on considère des objets de l'image. Pourquoi les limites d'une région représentant de la forêt devraient-elles dépendre de pixels représentant la



(A) Couleurs visibles (B) Partage des eaux

(C) Déplacement de la moyenne (D) Seuillage

FIGURE 2.14 – Comparaison visuelle de l'effet de différents algorithmes de segmentation appliqués à un extrait d'image satellitaire (A). Les paramètres des algorithmes de partage des eaux (B) de déplacement de la moyenne (C) n'ont pas été ajustés. Les régions produites sont difficiles à rapprocher des grands ensembles (eau, végétation, urbain) qui apparaissent sur l'image. Une méthode de segmentation par seuillage appliquée à la suite d'une préclassification des pixels de l'image d'origine (D) permet au contraire d'opérer une certaine interprétation sémantique de nombreuses zones, sans qu'il ait été nécessaire d'ajuster de quelconques paramètres. De plus, la taille des segments en (D) est beaucoup plus hétérogène qu'en (C) et surtout qu'en (B). Certaines méthodes de segmentation sont donc plus adaptées que d'autres à la délimitation de segments de tailles hétérogènes.

mer située à plusieurs dizaines de kilomètres ? Il ne semble y avoir aucune raison pour que les représentations de deux objets qui ne sont pas explicitement en relation sur le terrain s'influencent mutuellement du fait arbitraire d'un algorithme.

2.3.7.7 Description des segments

La délimitation de segments a une double conséquence :

Elle permet d'abord de donner une description de l'image plus synthétique : au lieu de prendre en compte la valeur spectrale, pseudo-spectrale ou texturale calculée en chaque pixel, il est possible de ne plus considérer qu'un résumé statistique sur les pixels de chaque segment (moyenne, variance, maximum...). La description de l'image est ainsi synthétisée par segment.

Ensuite, elle produit une *nouvelle information* qui est la forme des objets. Numériquement, la forme est décrite via des indices représentatifs comme l'élongation. Une recension de ces indices a été réalisée par Lehmann (2007).

2.3.7.8 Quelques réflexions sur les limites des objets

Nous avons précisé qu'un segment (*a fortiori* un objet de l'image) est doté d'un identifiant. Comme groupe de pixels, il est caractérisé par une frontière séparant les pixels qui font partie du segment de ceux qui n'en font pas partie. La définition de cette frontière mérite réflexion : a-t-elle une

épaisseur ? Fait-elle partie de l'entité[1] ? Est-elle floue (dans le sens de progressive), vague (dans le sens de nette mais incertaine) ? Casati et al. (1998) soulignent que la frontière est une marque d'individualité. Mais ces frontières, qui ont une place importante dans l'existence des entités qu'elles délimitent *existent*-elles vraiment ? Difficile de répondre...

Casati et al. (1998) (repris par Castilla and Hay (2008)) distinguent les discontinuités physiques (limites *bona fide*) des frontières conventionnelles (limites *fiat*), ce qui définit par extension deux types d'objets dans l'image, déterminés par le type de leurs limites et sachant qu'il existe des régions dont la frontière peut être composée de limites des deux types.

Quoi qu'il en soit, la définition des segments et de leurs frontières est une vaste question. L'obtention de segments sémantiquement cohérents nécessite que l'algorithme de segmentation bénéficie d'une connaissance experte... car un algorithme ne peut *deviner* ce que souhaiterait idéalement celui qui l'exécute. Cette connaissance prend diverses formes. Mais qu'il s'agisse d'ajuster un paramètre en fonction d'une expérience ou par essai/erreur ou qu'il s'agisse d'effectuer une préclassification préalable des pixels... c'est bien toujours la connaissance qui fait la qualité du résultat.

Enfin, le problème de la segmentation, même s'il est parfois rapproché de celui de la classification des entités, en diffère profondément. Quelle qu'elle soit, une classification ne crée pas de nouvelles entités mais se contente de regrouper et d'étiqueter des entités préexistantes. Au contraire, une segmentation donne naissance à des segments qui n'existaient pas avant l'opération. D'un point de vue ontologique, ces deux opérations n'ont donc rien à voir l'une avec l'autre.

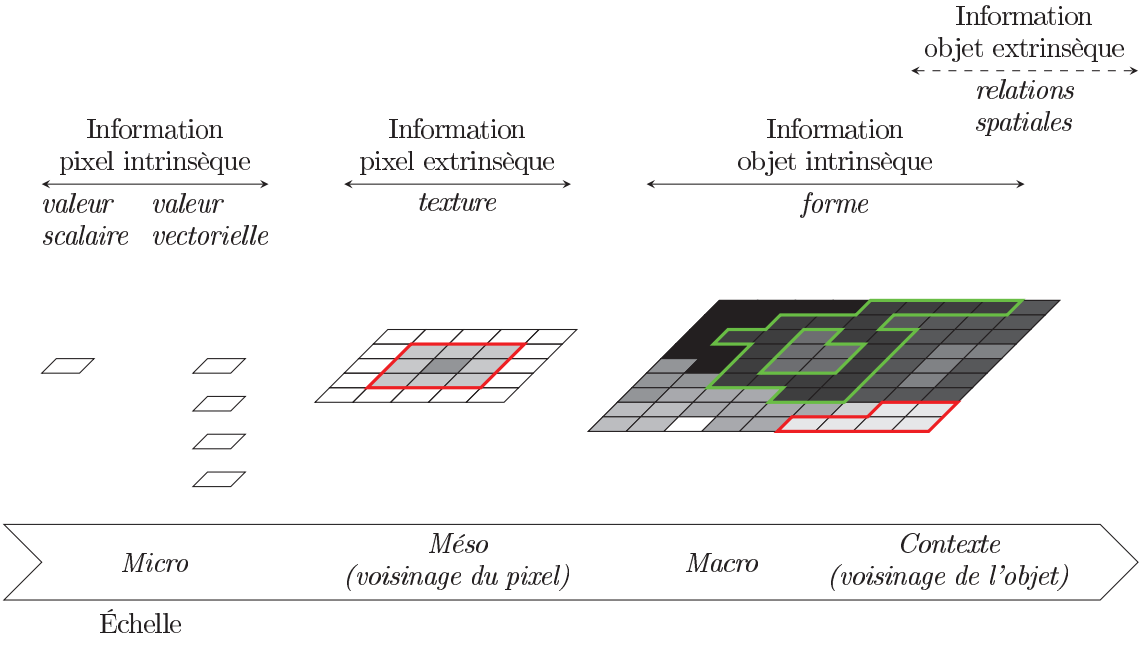


FIGURE 2.15 – Du pixel au contexte des objets dans l'image.

[1] Nous verrons dans la suite de ce document que la question des limites des entités n'est pas sans conséquence sur la représentation des relations spatiales quelles entretiennent les unes avec les autres. Elles jouent un double rôle. Premièrement, elles distinguent l'entité de ce qui lui est extérieur : un rôle d'isolement. Deuxièmement, elles contribuent à caractériser les relations entre l'entité et ce qui lui est extérieur : un rôle de mise en relation.

Pour résumer (Image et information)**.** *Castilla (2003) insiste sur le fait que l'information n'est pas extraite de l'image, mais produite par son analyse. Selon cette vision des choses, l'image est une simple source de données qui ne parle pas d'elle-même.*

*Les images satellitaires sont matricielles. Elles peuvent produire différentes sortes d'informations que l'on peut placer sur un axe indiquant l'échelle de traitement (*FIGURE *2.15).*

1. *Le* niveau du pixel *est atomique. Un pixel est porteur d'une valeur et d'une position dans l'image. Dans le cas d'une image multispectrale, un pixel est associé à plusieurs valeurs scalaires que l'on peut encore considérer comme une unique valeur vectorielle. Le pixel peut également être* porteur de sens *(longueur d'onde, grandeur physique...) à la lumière des métadonnées, en particulier à des fins de calibration.*

   Le calcul d'indices par le biais des opérations sur les images, permet de résumer en profondeur *une valeur vectorielle (plusieurs bandes spectrales) en une valeur scalaire (indice spectral). En un sens, ces traitements* appauvrissent les données *(remplacement d'une collection de valeurs scalaires par une seule) mais* enrichissent l'information *(l'indice spectral aide à l'interprétation de l'image).*

2. *Le calcul de textures requiert couramment de s'intéresser au* voisinage des pixels*. Le résultat est une valeur scalaire elle-même souvent affectée à un pixel et produit donc des images de texture dont les valeurs* résument en surface *la structure du voisinage des pixels.*

3. *La segmentation, quelle que soit la technique sur laquelle elle s'appuie, produit de nouvelles entités identifiables : les* segments d'image*. Les valeurs des pixels qui les composent offrent la possibilité de les décrire par toutes sortes de statistiques (spectrales, pseudo-spectrales ou texturales). La position des pixels confère également une information spatiale aux segments, mais ceux-ci disposent en plus d'une* forme *qui peut être décrite dans une certaine mesure (élongation...).*

4. *Enfin, les segments entretiennent des rapports avec leur voisinage (contexte spatial) que l'on peut caractériser, par exemple, à l'aide de relations topologiques spatiales.*

3

Formalisation et conceptualisation de la connaissance

Comme pour les domaines touchant aux images satellitaires, ce chapitre-ci s'applique à présenter l'état actuel de la représentation des connaissances en informatique. Le développement des paragraphes tente de faire ressortir la logique d'ensemble de l'évolution historique : attacher toujours plus de sémantique aux données.

Sommaire

3.1 Bref historique de la conceptualisation et de la représentation des connaissances **62**
3.1.1 Origine de l'ontologie . . . 62
3.1.2 Les ontologies en informatique . . . 63
3.2 Structuration et description des données . . . **63**
3.2.1 XML : l'ancêtre du web sémantique . . . 63
3.2.2 RDF : tout est description . . . 64
3.2.3 RDFS : la conceptualisation des ressources . . . 65
3.3 Les logiques de description . . . **67**
3.3.1 Définitions élémentaires . . . 67
3.3.2 Niveau assertionnel . . . 70
3.3.3 Niveau terminologique . . . 70
3.3.4 Notion d'interprétation . . . 71
3.3.5 Les constructeurs enrichissent l'expressivité d'un langage . . . 72
3.3.6 OWL : le web et l'expressivité des logiques de description . . . 74
3.3.7 Complexité des logiques de description . . . 75
3.3.8 Raisonneurs . . . 75
3.4 Représentation de la connaissance et enrichissement de l'information . . . **77**

3.1 Bref historique de la conceptualisation et de la représentation des connaissances

3.1.1 Origine de l'ontologie

3.1.1.1 Étude de l'être

En métaphysique, *l'ontologie* (au singulier) désigne l'étude de l'*être*. La question est très vaste et depuis les antiques philosophes hellènes Parménide, Héraclite et Gorgias, de nombreux penseurs s'y sont intéressés. D'après Victor COUSIN, « Platon est le génie de l'abstraction, Aristote celui de la classification » (Cousin, 1840). Nous héritons en Europe occidentale de cette pensée antique et sans nous poser trop de questions nous acceptons depuis l'enfance de découper le monde en *objets* qui seraient finalement des entités plus ou moins élémentaires. Puis, nous regroupons ces objets en *classes* en nous fondant plus ou moins arbitrairement sur des caractéristiques qu'ils partagent. Mais la classe n'est pas palpable comme le sont les objets.

Que ce soit sur les classes, les *concepts* ou les *idées*, le raisonnement sur des abstractions du monde est entré dans les mœurs depuis 2500 ans, d'abord repris par les Pères de l'Église, les scholastiques et les philosophes classiques. De nos jours encore, la question de l'être continue d'exciter la curiosité des épistémologues de tous domaines ; on peut citer à titre d'exemple le physicien Étienne KLEIN ou le psychanalyste Cornelius CASTORIADIS.

Il faut néanmoins distinguer différentes acceptions du terme, afin de ne pas transposer par mégarde une confusion en informatique. On peut pour cela s'inspirer de Verbeke (1952) :

- L'être des *individus* sensibles (ou *substances sensibles*). Par exemple, un homme ou un animal sont des individus, chacun doté de caractéristiques particulières.
- L'être des *abstractions* permettant de regrouper les individus (ou *substances immuables*). Les notions « d'homme » ou « d'animal » regroupent les individus de ces catégories et sont caractérisées par un ensemble de propriétés qui leurs sont propres. On peut rattacher cette acception de l'être au monde des *idées* platoniciennes.
- L'être des notions abstraites telles que le vrai, le beau, le juste, etc. Ces notions, bien que physiquement intangibles sont pourtant clairement désignées par la pensée, ce qui démontre leur existence[1] (ou *substances séparées*).

C'est là une vision des choses qui ne prétend pas à l'exclusivité. Elle peut être comparée dans une certaine mesure à la modélisation objet en informatique. Le premier sens de l'ontologie philosophique peut être rapproché de la vision des instances, le second de la vision des classes et le troisième n'est pas vraiment abordé.

3.1.1.2 Un système à base de connaissances

Le même terme d'« ontologie » utilisé aujourd'hui en sciences de l'information doit être à la fois considéré comme doté d'une signification propre, mais également comme l'héritier du sens élaboré par une riche histoire philosophique et épistémologique.

C'est ainsi que l'entend Saint-Dizier (2006), pour qui les ontologies en informatique sont un cas particulier de réseau sémantique dont les prémices ont été posés dès l'antiquité par Porphyre commentant Aristote. Les réseaux sémantiques offrent la possibilité de représenter des ensembles, des sous-ensembles et des éléments, ainsi que les relations qu'ils les relient deux à deux. En particulier, les réseaux sémantiques posent les bases de l'*instanciation* (un élément est membre d'un ensemble) et de la *subsomption* (un ensemble est sous-ensemble d'un autre).

Mais d'après la recension réalisée par Mariño Drews (1993), les graphes sémantiques eux-mêmes, dont l'origine revient à Quillian (1967), ne sont qu'un type de *système à base de connaissances* parmi d'autres. Cette famille englobe également les graphes conceptuels de Sowa (1991) élaborés pour représenter formellement des expressions du langage naturel. Elle englobe aussi les logiques descriptives dont nous reparlerons et les *frame-based languages* qui les ont précédées (Minsky, 1974).

[1] « Il faut donc tenir pour constant, Simmias, que si toutes ces choses, que nous avons toujours dans la bouche, existent véritablement, je veux dire le bon, le bien, et toutes les autres essences du même ordre... » (Platon, 1822)

3.1.2 Les ontologies en informatique

3.1.2.1 Une définition formelle

Dans les faits, en informatique, *les ontologies* (au pluriel) décrivent explicitement des concepts et les relations qu'ils entretiennent entre eux. On se limitera pourtant à une définition plus large et reconnue.

Définition 46 (Ontologie, d'après Gruber (1993))**.** Une *ontologie* est une spécification explicite d'une conceptualisation.

Cet énoncé minimal sert de référence, mais Gruber (1993) le propose dans un objectif de *partage* de la connaissance. Le partage de la connaissance est si fortement lié aux ontologies qu'elle se trouve parfois intégrée dans la définition (Studer et al., 1998). Et de fait, l'expression d'une conceptualisation a-t-elle une autre finalité que le partage ?

3.1.2.2 Partager une conceptualisation

Or, le web a rapidement joué un rôle important dans le partage de l'information. Si l'on constate cette communauté d'intérêt entre le web et les ontologies, il n'est pas surprenant dès lors que le développement des ontologies soit étroitement lié à l'évolution souhaitée du web vers un contenu sémantique (Gardiner et al., 2006).

Par *contenu sémantique* on entend que les éléments du contenu (les données) soient porteurs d'une information sur leur nature ou leur fonction[1]. Ontologies et métadonnées (voir la Définition 22) sont donc étroitement liées.

Mais l'enrichissement sémantique des données est progressif, comme l'illustre le développement des standards du W3C (*World Wide Web Consortium*).

Un point est sans doute essentiel : qu'une ontologie rencontre un large consensus ou une audience restreinte, elle vise toujours un certain partage. Cela peut-être un partage au sein d'une communauté humaine... mais également un partage entre l'homme et la machine.

3.2 Structuration et description des données

3.2.1 XML : l'ancêtre du web sémantique

L'apparition de XML (*Extensible Markup Language*) associé aux schémas XML constitue une première étape dans l'ajout de sémantique aux données textuelles (Horrocks et al., 2007).

Même si l'enrichissement sémantique n'est pas un objectif de XML (XML-a, 2008), il n'en reste pas moins qu'il intervient potentiellement à deux niveaux.

Premièrement, la sémantique découle de la structuration même du texte, indépendamment du nom des balises. L'imbrication des balises qui encadrent un texte peut être interprété, ce qui signifie qu'elle est déjà porteuse de sens et donne une information sur le texte lui-même (sur sa nature intrinsèque ou bien sur sa fonction dans l'ensemble des données).

Deuxièmement, XML ouvre la voie à un autre enrichissement sémantique des données en donnant des noms de balises et d'attributs (éventuellement compréhensibles par l'homme). XML gère l'homonymie des balises grâce au mécanisme des espaces de noms.

Les schémas XML sont une application du méta-langage[2] XML qui permet de définir un vocabulaire et une structure pour les applications XML.

[1] Par exemple, une chaîne de caractères doit être identifiée comme un *nom*, une *adresse* etc. Cette information, *au-delà des données* est une *métadonnée*.

[2] Un langage qui permet de construire d'autres langages.

3.2.2 RDF : tout est description

3.2.2.1 Fonction élémentaire de RDF

Le langage RDF (*Resource Description Framework*) est une recommandation du W3C (RDF-a, 2004) utilisée en particulier pour décrire les métadonnées relatives à une *ressource*.

Définition 47 (Ressource)**.** Le document (RDF-a, 2004) définit une *ressource* comme toute chose qui peut être référencée par une URI (*Uniform Resource Identifier*). La définition du vocabulaire (RDF-e, 2004) est plus large et précise que tout est ressource.

En RDF, toute description se rapporte à un *sujet* à l'aide de couples attribut-valeur. Dans le contexte RDF, les attributs sont appelés *prédicats* (ou propriétés) et leur valeur *objet*. Une description exprimée en RDF prend donc la forme d'un triplet sujet-prédicat-objet. Elle est relative au descripteur et doit être distinguée de la ressource elle-même. Rien n'empêche qu'une autre description de la même ressource existe par ailleurs et rien ne garantit la cohérence sémantique des différentes descriptions.

Remarque 22. Les couples RDF (sujet, objet) définissent un graphe de ressources et de valeurs littérales sur lesquelles les prédicats définissent des relations binaires.

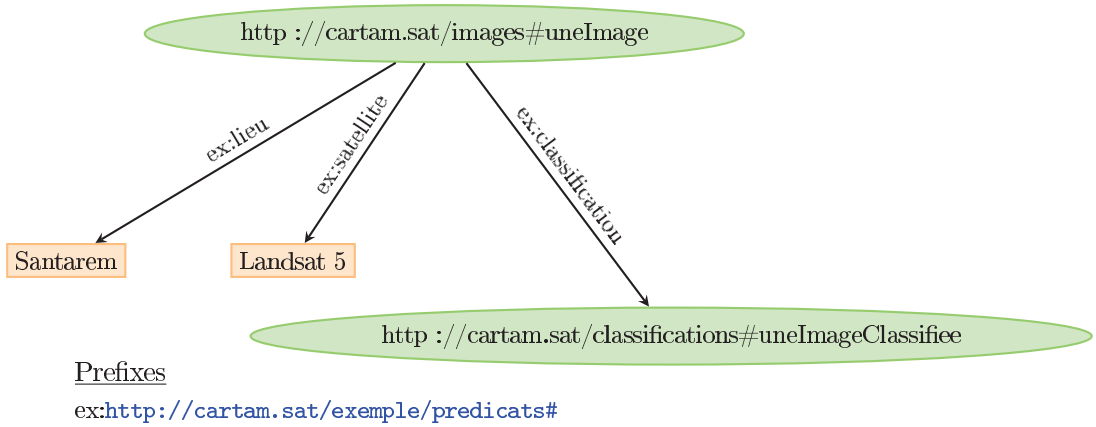


FIGURE 3.1 – Un exemple de graphe RDF.

RDF en tant que tel n'est pas lié à une syntaxe particulière. Toutefois, il se base sur les types de données définis par le langage XML-Schema et la spécification fournit également une application XML pour rendre les données RDF accessibles à la machine : RDF/XML (RDF-a, 2004). XML donne à la description des ressources une forme susceptible de bénéficier des avantages fournis par ce langage, à commencer par les parseurs et les interfaces de programmation. On peut donc écrire en syntaxe RDF/XML (CODE 3.1), l'équivalent de la représentation graphique précédente (FIGURE 3.1).

RDF peut également prendre d'autres formes, en utilisant des syntaxes notamment destinées à améliorer la lisibilité.

Comment lire cette description de ressource ? Les éléments structurels de RDF sont définis dans l'espace de noms identifié par l'URI `http://www.w3.org/1999/02/22-rdf-syntax-n#`. Dans le document XML, l'espace de nom des éléments est indiqué par des préfixes qui sont des alias plus faciles à manipuler et plus lisibles que les URI. C'est le cas pour l'élément racine *RDF* et pour l'élément *Description*.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:ex="http://cartam.sat/exemple/predicats#">
  <rdf:Description rdf:about="http://cartam.sat/images#uneImage">
    <ex:lieu>Santarem</ex:lieu>
    <ex:satellite>Landsat 5</ex:satellite>
    <ex:classification rdf:resource="http://cartam.sat/classifications#
      uneImageClassifiee"/>
  </rdf:Description>
</rdf:RDF>
```

CODE 3.1 – Un exemple de sérialisation en RDF/XML.

L'élément *Description* indique que le sujet (la ressource décrite) est identifié par l'URI `http://cartam.sat/images#uneImage` qui apparaît également dans la forme graphique. Ce sujet est décrit par trois prédicats définis dans l'espace de noms `http://cartam.sat/exemple/predicats#`.

Les prédicats *lieu* et *satellite* ont des valeurs littérales (ici des chaînes de caractères). Par contre, la valeur du prédicat *classification* fait référence à une ressource complexe référencée par une URI. Dans l'absolu, cette nouvelle ressource peut elle-même être décrite en RDF par d'autres couples (prédicat, objet). Une ressource vue comme un objet est aussi un sujet potentiel.

RDF est capable d'une part de définir des ressources et d'autre part de décrire des ressources préalablement définies en les référençant. En XML, ces deux fonctions se distinguent grâce aux attributs *ID* et *about*.

3.2.2.2 Réification

Dans RDF, les descriptions sont elles-mêmes vues comme des ressources que l'on peut décrire en les *réifiant*. La réification consiste en fait à donner corps à une relation. On spécifie alors explicitement le sujet, le prédicat et l'objet de la description réifiée. Dans ce cas, le prédicat *type*, nécessaire pour les réifications, désigne obligatoirement comme objet, la ressource RDF *Statement* (FIGURE 3.2).

La réification d'une relation permet de la décrire.

3.2.2.3 Premiers types de descriptions

En plus des valeurs littérales ou des ressources, RDF fournit des structures basiques pour considérer des ensembles de ressources de manière globale. Pour cela, RDF définit des conteneurs de différents *types* : pour des ensembles ordonnés (*Seq*) ou non ordonnés (*Bag*).

Le conteneur *Alt* indique une valeur possible parmi d'autres et lorsque tous les éléments de l'ensemble sont spécifiés de manière exhaustive, on utilise la collection *List*. Implicitement, ces conteneurs sont des descriptions typées.

3.2.3 RDFS : la conceptualisation des ressources

RDF utilise donc discrètement un typage des ressources qui font partie du langage. RDF-Schema complète RDF en définissant ces types formellement à partir de la notion de *Classe* et en utilisant lui-même les vocabulaires RDF et RDFS. Il permet également aux utilisateurs de RDF de définir leurs propres types de ressources et de propriétés. Les types de propriétés peuvent préciser les types de ressources acceptées comme sujet ou comme objet. Enfin, types de propriétés et types de ressources peuvent être organisés en hiérarchie en utilisant des prédicats de spécialisation. RDFS est donc un langage de description de vocabulaire qui enrichit la sémantique de RDF.

Un exemple[1] (CODE 3.2) déclare à l'aide de l'attribut *ID* plusieurs ressources qui sont des classes et des propriétés[2]. Suite à leur déclaration, d'autres descriptions peuvent se référer à ces

[1] Inspiré de `http://www.w3.org/TR/2000/CR-rdf-schema-20000327/`.

[2] Les types de ces descriptions sont RDFS *Class* ou RDF *Property*. De manière à alléger la syntaxe, on note usuellement la description d'une ressource de type « ns:ExempleDeClasse » par le nom de balise ns:ExempleDeClasse.

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]>
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xml:base="http://cartam.sat/exemple/satellites#">

<rdfs:Class rdf:ID="Satellite"/>
<rdfs:Class rdf:ID="SatelliteHeliosynchrone">
  <rdfs:subClassOf rdf:resource="#Satellite"/>
</rdfs:Class>
<rdfs:Class rdf:ID="SatelliteObservationTerre">
  <rdfs:subClassOf rdf:resource="#Satellite"/>
</rdfs:Class>
<rdfs:Class rdf:ID="SatelliteLandsat">
  <rdfs:subClassOf rdf:resource="#SatelliteHeliosynchrone"/>
  <rdfs:subClassOf rdf:resource="#SatelliteObservationTerre"/>
</rdfs:Class>
<rdfs:Class rdf:ID="CapteurMultispectral"/>

<rdf:Property rdf:ID="embarque">
  <rdfs:domain rdf:resource="#SatelliteObservationTerre"/>
  <rdfs:range rdf:resource="#CapteurMultispectral"/>
</rdf:Property>
<rdf:Property rdf:ID="bandes">
  <rdfs:domain rdf:resource="#CapteurMultispectral"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
</rdf:Property>
</rdf:RDF>
```

CODE 3.2 – Définition des types de ressources et de propriétés.

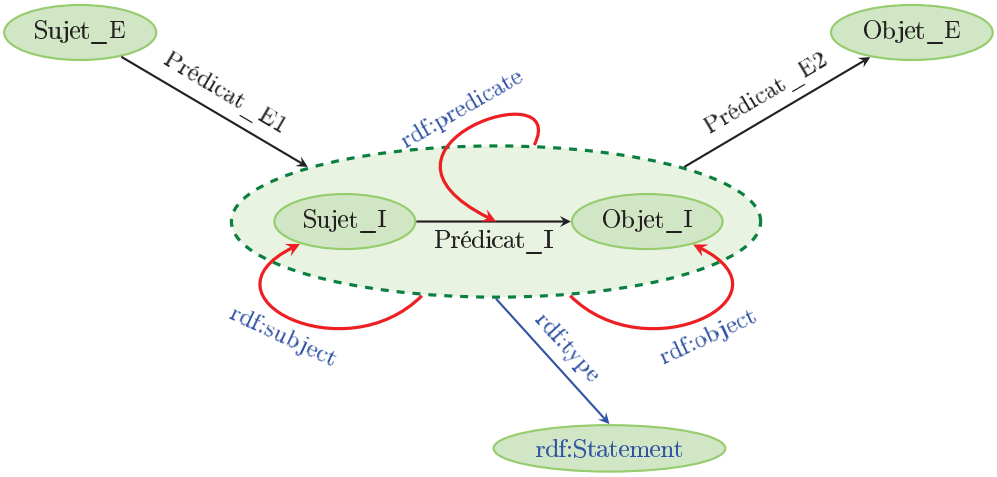


FIGURE 3.2 – Une illustration du mécanisme de réification dans RDF. Un triplet peut être vu comme une ressource et jouer le rôle d'un objet ou d'un sujet dans un des triplets. Les prédicats, représentés par des flèches, sont de trois sortes. Ceux de couleur noire indiquent que le triplet réifié, comme ressource, peut jouer le rôle d'un sujet ou d'un objet dans d'autres triplets. Le prédicat de couleur bleue représente explicitement le typage de la ressource constituée par le triplet. Enfin, les prédicats de couleur rouge indiquent le sujet, le prédicat et l'objet du triplet réifié.

ressources à l'aide de l'attribut *about*. La référence à ces ressources requiert en général leur URI absolue composée de l'identifiant de la ressource précédé du caractère # concaténé à l'URI de base[1].

L'exemple suivant (CODE 3.3) utilise des instances de ces classes et propriétés, c'est-à-dire la description de nouvelles ressources typées à l'aide des ressources précédentes.

3.3 Les logiques de description

3.3.1 Définitions élémentaires

Définition 48 (Logique de description, d'après Baader and Nutt (2003); Baader et al. (2007))**.** Une *logique de description* est un langage formel de représentation des connaissances doté d'une sémantique formelle basée sur la logique.

De manière générale, on considère que les logiques de description sont *décidables* (Ben Mohamed, 2010).

Définition 49 (Décidabilité (algorithmique), d'après Bournez et al. (2010))**.** Une famille d'énoncés vrais est décidable si elle forme un ensemble dont on peut décider par un algorithme si un énoncé donné y figure ou pas.

[1]L'URI de base est donnée par l'attribut xml:base du document, ou en son absence, par l'URI du document courant.

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:ex="http://cartam.sat/exemple/satellites#"
         xml:base="http://cartam.sat/exemple/unSatellite#">

  <ex:SatelliteLandsat rdf:ID="Landsat5">
    <ex:embarque rdf:resource="#ThematicMapper"/>
  </ex:SatelliteLandsat>
  <ex:CapteurMultispectral rdf:ID="ThematicMapper">
    <ex:bandes rdf:datatype="&xsd;integer">7</ex:bandes>
  </ex:CapteurMultispectral>
</rdf:RDF>
```

CODE 3.3 – Utilisation des types de ressources et de propriétés.

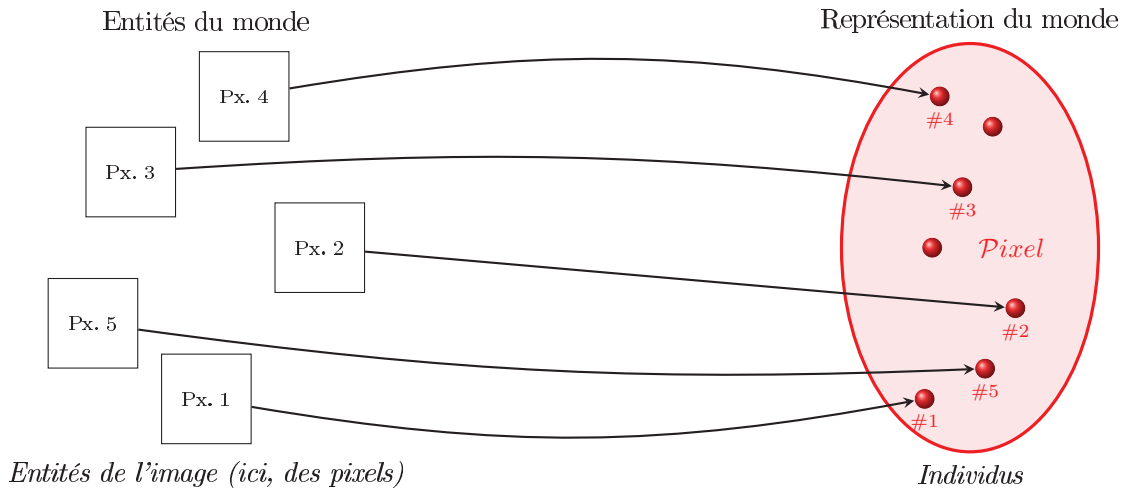


FIGURE 3.3 – La représentation du monde par les logiques de description. Des entités du monde telles que des pixels (à gauche) sont représentées par des individus (à droite) généralement dotés d'un identifiant (#i). Les individus peuvent appartenir à des ensembles : ici, le concept de *Pixel*.

Définition 50 (Individu)**.** On appelle *individu* la représentation d'une entité du monde (FIGURE 3.3).

Définition 51 (Concept, d'après Nardi and Brachman (2003))**.** Un *concept* représente un ensemble d'individus (FIGURE 3.3). Il peut être vu comme un prédicat unaire : $Pixel(\#1)$ signifie que l'individu #1 est membre du concept *Pixel*.

Remarque 23 (Notation des concepts)*.* On note généralement $\top$ le concept le plus général, $\bot$ le concept le plus spécifique.

Définition 52 (Concepts atomique, composé, primitif et défini, d'après Baader and Nutt (2003))**.** Les logiques de description représentent un concept par un symbole (*concept atomique*) ou une combinaison de symboles (*concept composé*). Un concept atomique peut être *défini* par une combinaison de concepts plus élémentaires (définition par intention) ou par énumération de ses membres (définition par extension). Dans le cas contraire, c'est un *concept primitif*.

Définition 53 (Relation binaire, d'après Deschamps et al. (1999))**.** Soit E un ensemble quelconque. Si G est une partie de $E \times E$, le prédicat $\mathcal{R}$ défini par :

$$\mathcal{R}(x, y) \iff (x, y) \in G$$

est appelé *relation binaire* sur E. On note la plupart du temps $x\mathcal{R}y$.

Définition 54 (Rôle, d'après Nardi and Brachman (2003))**.** Un *rôle* représente un ensemble de paires d'individus (FIGURE 3.4). Il peut être vu comme un prédicat binaire : $Adjacent(\#1, \#2)$ signifie que l'individu $\#1$ est en relation avec l'individu $\#2$ par le rôle $Adjacent$.

Remarque 24 (Notation des rôles)**.** On note généralement $\top_{\mathcal{R}}$ le rôle le plus général et $\bot_{\mathcal{R}}$ le rôle le plus spécifique.

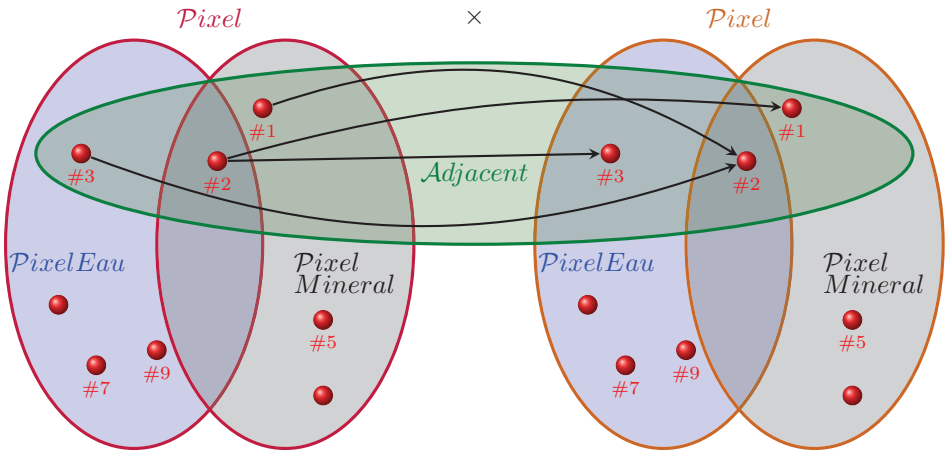


FIGURE 3.4 – Les relations entre individus sont des couples d'individus. Ici, les individus représentent quelques pixels de différentes sortes. Une relation d'adjacence lie un certain nombre d'individus entre eux. L'individu $\#1$, par exemple, est adjacent au $\#2$. L'ensemble de ces couples définit le rôle $Adjacent$ sur $Pixel \times Pixel$.

Les logiques de description permettent de mettre en place des *bases de connaissance* et de raisonner sur leur contenu (Baader and Nutt, 2003). Les entités du monde sont représentées par des *individus*. Une base de connaissances est constituée d'un niveau *terminologique* et d'un niveau *assertionnel*.

3.3.2 Niveau assertionnel

La *ABox* représente le monde par des assertions d'appartenance d'un individu à un concept et de relation entre deux individus. C'est une description de la connaissance plutôt extensionnelle qui relève d'un problème particulier (Nardi and Brachman, 2003).

Remarque 25 (Axiomes assertionnels (ou assertions) d'appartenance et de relation). Soit C un concept, R un rôle et a et b deux individus. On note :

- $C(a)$ ou $a : C$ l'assertion selon laquelle a appartient à C ;
- $R(a, b)$ ou $(a, b) : R$ l'assertion selon laquelle a est lié à b par le rôle R.

Pour résumer (Abox). *Une ABox est un ensemble fini d'assertions (Baader et al., 2007) sur des individus définissant une représentation du monde.*

3.3.3 Niveau terminologique

La *TBox* définit une terminologie (déclaration de concepts et de rôles) et la taxinomie (leur organisation hiérarchique). C'est une description de la connaissance plutôt intentionnelle qui traite d'un domaine en général (Nardi and Brachman, 2003).

La TABLE 3.1 montre les deux types d'axiome terminologique présents dans la TBox (Espinasse, 2009).

Axiome terminologique	Notation	Sémantique
Subsomption	$C_1 \sqsubseteq C_2$	Les instances de C_1 sont des instances de C_2.
Équivalence	$C_1 \equiv C_2$	Les instances de C_1 sont les instances de C_2.

TABLE 3.1 – Les axiomes terminologiques (C_1 et C_2 sont des concepts).

L'axiome d'équivalence dont le membre de gauche est un concept atomique est également appelé *définition*[1]. On remarquera que la subsomption (ou *spécialisation*) est en fait l'expression d'une condition *nécessaire*, alors que l'équivalence exprime une condition *nécessaire et suffisante*. Il en ressort naturellement que l'équivalence peut être ramenée à une double subsomption (Nardi and Brachman, 2003).

Remarque 26 (D'après Nardi and Brachman (2003)). Pour une terminologie donnée,

1. pour un nom donné, un concept est défini au plus une fois ;
2. une définition est *acyclique*, c'est-à-dire que dans une définition donnée, le concept défini dans le membre de gauche de l'équivalence ne peut être directement ou indirectement contenu dans le membre de droite.

[1] On peut représenter la définition par la notation suivante : $C_1 \doteq C_2$ (où C_1 et C_2 sont des concepts).

Remarque 27 (Concept et expression symbolique). Dans les définitions qui ont été adoptées, le mot *concept* se distingue de son acception courante qui n'en retient somme toute que la définition extensionnelle : un concept se trouve défini par les entités qui s'y réfèrent et sa représentation symbolique n'a pas d'importance.

Dans les logiques de description, les concepts sont plus étroitement liés à leur expression symbolique. Ainsi, considérons la définition suivante (C_1, C_2 et C_3 sont des concepts atomiques) :

$$C_1 \equiv C_2 \sqcap C_3$$

On dirait couramment que C_1 d'une part et $C_2 \sqcap C_3$ d'autre part renvoient au *même concept* car leurs extensions sont identiques. Dans le cadre des définitions formelles proposées, ces deux concepts sont équivalents mais différents (remarquons au passage que C_1 est atomique et défini, alors que $C_2 \sqcap C_3$ est composé).

Pour résumer (TBox). *Une TBox est un ensemble d'axiomes terminologiques sur des concepts et des rôles. Certains axiomes définissent des concepts ou des rôles élaborés à l'aide de concepts et de rôles plus simples articulés par des constructeurs.*

Au niveau terminologique, certains travaux limitent la TBox aux axiomes sur les concepts et définissent une *RBox* contenant les axiomes sur les rôles (Pan, 2004; Kazakov, 2008; Rudolph, 2011). Les axiomes de la RBox, ou bien sont de la même forme que les axiomes relatifs aux concepts (prédicats binaires de subsomption et d'équivalence en TABLE 3.1), ou bien prennent la forme de prédicats unaires s'il s'agit de définir par exemple la transitivité ou la fonctionnalité d'un rôle.

3.3.4 Notion d'interprétation

Comme langage, une logique de description manipule des signifiants. Pour décrire la sémantique d'une logique de description, il faut l'interpréter en reliant les signifiants à des signifiés. L'*interprétation* $\mathcal{I}$ est donc composée :

- d'un *domaine d'interprétation* constitué par l'ensemble des éléments signifiés, noté $\Delta^{\mathcal{I}}$;
- et d'une *fonction d'interprétation* notée $.^{\mathcal{I}}$.

La fonction d'interprétation $.^{\mathcal{I}}$ a une triple utilité en associant (Baader and Nutt, 2003) :

- à chaque individu a, un élément $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$;
- à chaque concept atomique A, un ensemble $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$;
- à chaque rôle atomique R, une relation binaire $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

Définition 55 (Base de connaissances, d'après Baader et al. (2007)). Une *base de connaissances* $\mathcal{K}$ est un couple $(\mathcal{T}, \mathcal{A})$ où $\mathcal{T}$ est une TBox et $\mathcal{A}$ une ABox.

Définition 56 (Modèle, d'après Baader et al. (2007))**.** C_i des concepts, a_i des individus, R un rôle, $\mathcal{T}$ une TBox, $\mathcal{A}$ une ABox, $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ une base de connaissances et $\mathcal{I}$ une interprétation.

$\mathcal{I}$ est un *modèle* de l'axiome d'inclusion $C_1 \sqsubseteq C_2$ si $C_1^{\mathcal{I}} \subseteq C_2^{\mathcal{I}}$.

$\mathcal{I}$ est un *modèle* de $\mathcal{T}$ ($\mathcal{I} \models \mathcal{T}$) si elle est un modèle de tous les axiomes d'inclusion de $\mathcal{T}$.

$\mathcal{I}$ est un *modèle* de l'assertion $R(a_1, a_2)$ si $R^{\mathcal{I}}(a_1^{\mathcal{I}}, a_2^{\mathcal{I}})$.

$\mathcal{I}$ est un *modèle* de l'assertion $C(a)$ si $C^{\mathcal{I}}(a^{\mathcal{I}})$.

$\mathcal{I}$ est un *modèle* de $\mathcal{A}$ ($\mathcal{I} \models \mathcal{A}$) si elle est un modèle de toutes les assertions de $\mathcal{A}$.

$\mathcal{I}$ est un *modèle* de $\mathcal{K}$ ($\mathcal{I} \models \mathcal{K}$) si elle est un modèle de $\mathcal{T}$ et de $\mathcal{A}$.

Si $\mathcal{I}$ est un modèle de la formule axiomatique ou assertionnelle ϕ, on note $\mathcal{I} \models \phi$.

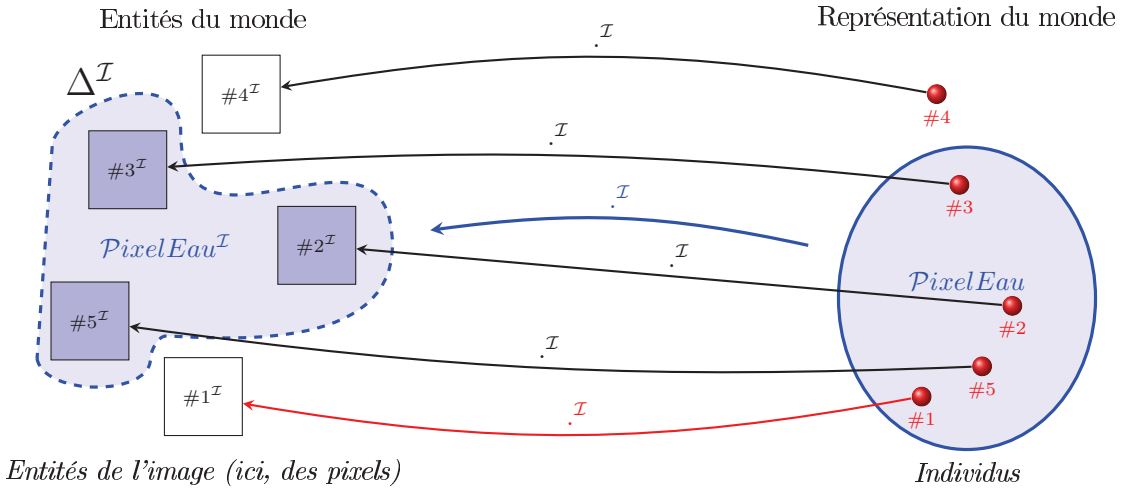


FIGURE 3.5 – Une fonction d'interprétation $\cdot^{\mathcal{I}}$ associe à chaque individu, un élément du domaine d'interprétation $\Delta^{\mathcal{I}}$ (ici, des pixels). Toutefois, dans cette illustration, l'interprétation $(\cdot^{\mathcal{I}}, \Delta^{\mathcal{I}})$ ne constitue pas un modèle : on a en effet $(PixelEau(\#1) \wedge \neg PixelEau^{\mathcal{I}}(\#1^{\mathcal{I}}))$.

Définition 57 (Consistance, d'après Baader et al. (2007))**.** Une base de connaissances est *consistante* si elle a un modèle.

Définition 58 (Satisfaisabilité, d'après Baader and Nutt (2003))**.** Un concept C est *satisfaisable* relativement à une terminologie $\mathcal{T}$ si il existe un modèle $\mathcal{I}$ de $\mathcal{T}$ tel que $C^{\mathcal{I}} \neq \emptyset$. On dit alors que $\mathcal{I}$ est un *modèle* de C.

3.3.5 Les constructeurs enrichissent l'expressivité d'un langage

Définition 59 (Expressivité)**.** L'expressivité d'une logique de description désigne sa capacité à construire une terminologie capable de décrire des faits du monde.

Plus l'expressivité d'une logique de description est élevée, plus les terminologies qui l'utilisent ont un potentiel important de représentation des phénomènes. L'expressivité dépend directement

de *constructeurs* grâce auxquels des concepts primitifs peuvent être associés pour donner naissance à des concepts définis.

Les formes d'axiomes terminologiques sur les rôles sont également considérées comme des constructeurs (Baader et al., 1990).

3.3.5.1 La sémantique de base

Les langages $\mathcal{FL}$ et $\mathcal{FL}^-$ Brachman and Levesque (1984) proposèrent d'abord une logique de description basée sur les langages de *frame*. Cette logique de description notée $\mathcal{FL}$ (pour *Frame-based Language*) comprend les constructeurs indiqués en TABLE 3.2. On note $\mathcal{FL}^-$ le langage $\mathcal{FL}$ privé de la restriction de rôle.

Constructeur	Notation	Interprétation sémantique
Conjonction	$C_1 \sqcap C_2$	$C_1^{\mathcal{I}} \cap C_2^{\mathcal{I}}$
Quantification universelle complète	$\forall R.C$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \forall c^{\mathcal{I}} \in \Delta^{\mathcal{I}}, R^{\mathcal{I}}(a^{\mathcal{I}}, c^{\mathcal{I}}) \Rightarrow c^{\mathcal{I}} \in C^{\mathcal{I}})\}$
Quantification existentielle limitée	$\exists R.\top$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \exists b^{\mathcal{I}} \in \Delta^{\mathcal{I}} : R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}))\}$
Restriction de rôle	$R.C$	$\{(a^{\mathcal{I}}, c^{\mathcal{I}}) \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(a^{\mathcal{I}}, c^{\mathcal{I}}) \wedge c^{\mathcal{I}} \in C^{\mathcal{I}}\}$

TABLE 3.2 – Les constructeurs de la logique $\mathcal{FL}$ (C_i est un concept, R un rôle).

Le langage $\mathcal{AL}$ et les constructeurs $\mathcal{E}$, $\mathcal{U}$ et $\mathcal{C}$ Schmidt-Schauss and Smolka (1991) ont ensuite établi une logique de description minimale $\mathcal{AL}$ (pour *Attributive Language*) qui revient à ajouter la négation atomique au langage $\mathcal{FL}^-$.

Constructeur	Notation	Interprétation sémantique
Négation atomique	$\neg C$	$\Delta^{\mathcal{I}} \backslash C^{\mathcal{I}}$

TABLE 3.3 – La négation atomique $\mathcal{AL}$ (C est un concept atomique).

Ils proposèrent également d'autres constructeurs (TABLE 3.4) permettant d'enrichir la sémantique du langage $\mathcal{AL}$ quoique de manière parfois redondante ($\mathcal{ALC}$ et $\mathcal{ALUE}$ sont équivalents). Parallèlement aux constructeurs de concepts, on peut également définir des constructeurs de rôles autorisant par exemple leur transitivité. On abrège généralement par la notation $\mathcal{S}$ les langages équivalents à $\mathcal{ALCR}^+$ (Fournier-Viger, 2005; Baader et al., 2007).

	Constructeur	Notation	Interprétation sémantique
$\mathcal{C}$	Négation non-atomique	$\neg C$	$\Delta^{\mathcal{I}} \backslash C^{\mathcal{I}}$
$\mathcal{U}$	Disjonction	$C_1 \sqcup C_2$	$C_1^{\mathcal{I}} \cup C_2^{\mathcal{I}}$
$\mathcal{E}$	Quantification existentielle complète	$\exists R.C$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \exists c^{\mathcal{I}} \in \Delta^{\mathcal{I}} : (R^{\mathcal{I}}(a^{\mathcal{I}}, c^{\mathcal{I}}) \wedge c^{\mathcal{I}} \in C^{\mathcal{I}})\}$
$\mathcal{R}^+$ ou $\mathcal{S}$	Transitivité des rôles	R^+	

TABLE 3.4 – De $\mathcal{AL}$ à $\mathcal{S}$ (C_i est un concept, R un rôle).

3.3.5.2 Des constructeurs supplémentaires, domaines concrets

Des constructeurs logiques supplémentaires sont utilisés pour accroître l'expressivité des langages basée sur les logiques de description (TABLE 3.5).

[1] Les rôles sont atomiques.

	Constructeur	Notation	Interprétation sémantique
$\mathcal{F}$	Fonctionnalité	$\leq 1R$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \text{card}(\{b^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}})\}) \leq 1\}$
$\mathcal{N}$	Restriction de cardinalité	$\leq nR$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \text{card}(\{b^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}})\}) \leq n\}$
		$\geq nR$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \text{card}(\{b^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}})\}) \geq n\}$
$\mathcal{Q}$	Restriction de cardinalité qualifiée	$\leq nR.C$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \text{card}(\{c^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(a^{\mathcal{I}}, c^{\mathcal{I}}) \wedge c^{\mathcal{I}} \in C^{\mathcal{I}}\}) \leq n\}$
		$\geq nR.C$	$\{a^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid \text{card}(\{c^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(a^{\mathcal{I}}, c^{\mathcal{I}}) \wedge c^{\mathcal{I}} \in C^{\mathcal{I}}\}) \geq n\}$
$\mathcal{O}$	Énumération	$\{a_1 \dots a_n\}$	$\{a_1^{\mathcal{I}} \dots a_n^{\mathcal{I}}\}$
$\mathcal{B}$	Role filler	$\exists R.\{a\}$	$\{c^{\mathcal{I}} \in \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(c^{\mathcal{I}}, a^{\mathcal{I}})\}$
$\mathcal{H}$	Hiérarchie de rôles atomiques[1]	$R_1 \sqsubseteq R_2$	$R_1^{\mathcal{I}} \subseteq R_2^{\mathcal{I}}$
$\mathcal{R}$	Hiérarchie de rôles non-atomiques[2]	$R_1 \sqsubseteq R_2$	$R_1^{\mathcal{I}} \subseteq R_2^{\mathcal{I}}$
$\mathcal{I}$	Inversion de rôle	R^{-1}	$\{(c_1^{\mathcal{I}}, c_2^{\mathcal{I}}) \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid R^{\mathcal{I}}(c_2^{\mathcal{I}}, c_1^{\mathcal{I}})\}$

TABLE 3.5 – Constructeurs courants (C_i sont des concepts, R_i des rôles).

Les constructeurs courants permettent d'utiliser des concepts, des rôles... et même des individus. Mais la représentation de la connaissance a également rendu nécessaire la représentation de types de données élémentaires par le biais des *domaines concrets*.

Définition 60 (Domaine concret, d'après Baader and Sattler (1998))**.** On peut étendre les logiques de description avec un *domaine concret* (ou un type de donnée), c'est-à-dire des ensembles concrets d'entiers ou de chaînes de caractères. On peut également utiliser des prédicats définis sur ces ensembles, tels que les relations d'ordre usuelles sur les nombres. On indice habituellement par $_{(\mathcal{D})}$ l'extension d'une logique de description par un domaine concret. Il est également possible d'intégrer des fonctions d'agrégation sur les domaines concrets, telles que les calculs de somme, de moyenne...

3.3.6 OWL : le web et l'expressivité des logiques de description

RDFS permet de hiérarchiser les classes et les propriétés. OWL (*Web Ontology Language*) permet en plus de construire des classes complexes par combinaison de classes de base. Pour cela, OWL se base sur les constructeurs des logiques de description.

La première version, OWL 1 se décline en trois sous-langages qui correspondent chacun à une logique de description. Ce sont, par ordre croissant d'expressivité (Horrocks et al., 2007) :

1. OWL-Lite, doté de la logique de description $\mathcal{SHIF}_{(\mathcal{D})}$.
2. OWL-DL, qui permet d'utiliser les constructeurs correspondant au langage $\mathcal{SHOIN}_{(\mathcal{D})}$;
3. OWL-Full sort du cadre des logiques de description et n'est pas décidable.

OWL 1 permet donc la structuration de conceptualisations en offrant la possibilité d'utiliser des constructeurs logiques mais sans restriction concernant la décidabilité au-delà de OWL-DL.

La version OWL 2 (OWL-a, 2012) s'appuie sur le langage $\mathcal{SROIQ}_{(\mathcal{D})}$ (OWL-d, 2012) pour définir trois profils décidables d'expressivité croissante : OWL-EL, OWL-QL et OWL-RL.

Dans OWL, une classe est à la fois une instance et une sous-classe de classe RDFS. Les classes d'OWL sont utilisées pour représenter les *concepts* de la pensée humaine (OWL-h, 2012) et sont assimilées aux concepts des logiques de description. De même, les propriétés OWL correspondent aux rôles des logiques de description.

Les domaines concrets ont été pris en compte dès la première version d'OWL sans toutefois permettre de représenter les intervalles de valeurs, les relations entre les valeurs attachées à un

[2]Les rôles ne sont pas atomiques mais le constructeur impose une *restriction de régularité* qui revient à le définir comme acyclique de manière à préserver la décidabilité du langage (Grau et al., 2008).

objet ou à des objets différents ainsi que les fonctions agrégatives. OWL 2 n'apporte qu'une réponse partielle à la représentation des intervalles par restriction de types XSD (Grau et al., 2008). Les fonctions agrégatives sont l'objet de recherches depuis une quinzaine d'années au moins (Baader and Sattler, 1998, 2003), mais elles n'apparaissent pas pour autant dans les standards.

3.3.7 Complexité des logiques de description

Même si des considérations plus fines peuvent tempérer ce constat global (Baader, 1996), il faut s'attendre à ce que la complexité algorithmique nécessaire à la résolution des problèmes d'inférence augmente avec l'expressivité d'une logique de description. Pour un algorithme informatique, la complexité fait référence à la progression du temps de calcul (complexité en temps) ou de l'espace mémoire nécessaire (complexité en espace) en fonction d'un paramètre de l'algorithme.

Il est parfois possible de cerner la complexité intrinsèque d'un problème en l'encadrant par une borne inférieure et une borne supérieure. L'identité de ces deux complexités décrit la complexité intrinsèque du problème en fonction du paramètre considéré (Bournez et al., 2010).

Si l'on se rapporte à cette théorie, les logiques de descriptions peuvent être associées à des classes de complexité. La TABLE 3.6 indique la classe de quelques logiques descriptives courantes. Pour d'autres, le problème reste ouvert.

Logique desc.	Équiv. en OWL	Classe	Référence
$\mathcal{SHIF}_{(\mathcal{D})}$	OWL 1 Lite	ExpTime	Horrocks et al. (2007); Bock et al. (2008)
$\mathcal{SHOIN}_{(\mathcal{D})}$	OWL 1 DL	NExpTime	Horrocks et al. (2007); Bock et al. (2008)
$\mathcal{SHIQ}$	—	ExpTime	Tsarkov and Horrocks (2006)
$\mathcal{SHOIQ}$	—	NExpTime	Tsarkov and Horrocks (2006)
$\mathcal{SROIQ}$	OWL 2	N2ExpTime	Grau et al. (2008); Kazakov (2008)

TABLE 3.6 – Quelques logiques descriptives et leur classe de complexité.

Il ressort de ce premier examen que l'utilisation des logiques de description à des fins d'inférence présente le risque d'une évolution très coûteuse, ce qui est susceptible de poser des problèmes (ou au moins des limites) dans les applications pratiques, malgré l'intérêt de l'approche.

3.3.8 Raisonneurs

Un raisonneur sur les logiques de description est un logiciel capable d'inférer de la connaissance en se basant sur des règles logiques. L'information supplémentaire ne sort pas de nulle part et le fonctionnement d'un raisonneur (FIGURE 3.6 et FIGURE 3.7) s'inscrit dans le schéma d'enrichissement de l'information décrit plus haut (FIGURE 2.6 page 43). Le processus d'inférence opéré par un raisonneur se décompose en plusieurs problèmes (Baader and Hollunder, 1991; Bock et al., 2008). On distinguera toutefois deux questions principales : la classification et la réalisation.

Définition 61 (Classification d'une base de connaissances, d'après Bock et al. (2008))**.** La *classification* d'une base de connaissances consiste à compléter la hiérarchie explicite des concepts par la hiérarchie implicite.

Définition 62 (Réalisation d'une base de connaissances, d'après Bock et al. (2008))**.** La *réalisation* d'une base de connaissances consiste à calculer le concept le plus spécifique auquel appartient chaque instance.

Remarque 28 (Réalisation et classification). Dans une certaine mesure, la réalisation peut être vue comme une classification au sens large, puisqu'elle consiste à expliciter des groupes d'éléments.

La mécanique interne d'un raisonneur est difficile à appréhender si l'on n'est pas directement concerné par le développement actif d'un logiciel de ce type (Bail et al., 2010). On peut voir un raisonneur comme un outil strictement dédié à l'aspect de *traitement* d'un problème qui prend en paramètre une *connaissance* sur la terminologie et les faits relatifs au problème. Un raisonneur est donc très générique et ne dépend pas du domaine d'un problème, mais du langage (la logique de description) dans lequel est exprimée la connaissance qu'il peut prendre en charge. Ce type d'outil démontre l'intérêt de la séparation de la connaissance et du traitement : la modification ou la substitution de la représentation d'un problème n'a pas d'impact en terme de modification algorithmique, pourvu que la logique de description ne change pas.

L'essor et les standards du web ont encouragé les développeurs (Gardiner et al., 2006). Les raisonneurs sont nombreux[1], mais cette diversité n'implique pas forcément un vaste choix d'applications.

Tout d'abord, les raisonneurs populaires sont pour la plupart accessibles via une interface de programmation en Java (Hermit, Pellet, QuOnto, Kaon2). D'autre part, il est courant que la sérialisation de la connaissance suive une syntaxe fonctionnelle proche du LISP. Le support des standards OWL et RDF/XML du W3C n'est assuré que via l'interface de programmation OWL-API, elle aussi en Java ou bien en utilisant une interface DIG (*DL Implementation Group*), système sur lequel il est assez difficile de se documenter[2] (FaCT++, RacerPro). Enfin, si l'on considère des contraintes budgétaires (gratuité), éthiques (licence libre) et de collaboration scientifique (maintenance par le développeur), les rangs des candidats s'éclaircissent rapidement.

L'éditeur d'ontologies *Protégé*[3] propose l'accès à des raisonneurs de manière à vérifier la consistance de la connaissance ou à classer des individus. Mais l'intégration d'un raisonneur dans une chaîne de traitements qui ne serait pas implémentée en Java n'est pas triviale si l'on veut également utiliser les standards du W3C.

Dépendant en particulier de la complexité des logiques de description, ils sont confrontés à un important problème de passage à l'échelle (Weithöner et al., 2006). Des travaux de parallélisation des algorithmes permettent d'espérer un certain amortissement du temps de calcul (Wu and Haarslev, 2012) de la logique descriptive ALC dont la complexité est polynomiale en espace[4], mais cela ne résout pas le problème des complexités exponentielles induites par des logiques plus expressives.

De nombreux raisonneurs implémentent un algorithme dit « des tableaux », comme c'est le cas pour FaCT++, mais cette méthode n'est pas exclusive : Hermit, Kaon2, QuOnto implémentent des méthodes différentes. De plus, les raisonneurs implémentent des optimisations algorithmiques (Tsarkov and Horrocks, 2006; Bail et al., 2010). Chaque raisonneur a donc ses caractéristiques qui le rendent plus efficace que ses concurrents pour certains problèmes et moins efficace pour d'autres. En particulier, certains supportent mieux le passage à l'échelle sur le nombre de concepts et d'autres sur le nombre d'individus. Outre les différences techniques citées plus haut, cet aspect peut également jouer un rôle dans le choix d'un raisonneur adapté (Bock et al., 2008).

Mais la considération du temps de calcul n'a de sens que si la réponse est correcte !

Malgré leur intérêt considérable, les raisonneurs ne sont pas des outils de vérité absolue. Plusieurs études ont été consacrées à leur comparaison. Et c'est là un autre vaste problème des raisonneurs, puisque les développeurs testent les algorithmes sur de petits exemples aisément compréhensibles. Pour ce qui est de la réponse sur des exemples plus volumineux, c'est la comparaison des résultats donnés par différents raisonneurs qui fait foi, mais ceux-ci ne concordent pas toujours (Gardiner et al., 2006) !

[1] La page http://www.cs.man.ac.uk/~sattler/reasoners.html propose une recension intéressante.

[2] Le système d'interfaçage DIG est basé sur le respect d'un schéma XML, mais l'interface de programmation utilisée par les raisonneurs RacerPro (écrit en LISP) et FaCT++ (écrit en C++) est elle aussi en Java : http://dig.sourceforge.net/javadoc-reasoners/1.1/

[3] http://protege.stanford.edu/. *Protégé* est un logiciel libre sous Licence Publique Mozilla 1.1 (http://www.mozilla.org/MPL/1.1/).

[4] Voir Schmidt-Schauss and Smolka (1991) pour la borne inférieure de la satisfaisabilité d'un concept et Tobies (2001) pour la borne supérieure. Voir Baader et al. (2005) pour la borne supérieure de la consistance des faits.

Enfin, un raisonneur repose sur l'*hypothèse du monde ouvert* (Bail et al., 2010).

Définition 63 (Hypothèse du monde ouvert, d'après Hustadt (1994))**.** Un fait qui n'est pas contenu dans une base de connaissances peut être vrai.

La description des faits doit être adaptée à ce cadre, puisqu'il ne suffit pas de ne pas le décrire pour le considérer comme faux. Les résultats inférés par un raisonneur doivent être considérés en conséquence (Falomir et al., 2011).

Nous aurons l'occasion d'observer plus en détail l'impact concret de cette hypothèse dans le développement de l'approche. Il faut tout de même prendre en compte qu'elle complique la représentation des connaissances et les tâches d'implémentation (Doherty et al., 2000).

3.4 Représentation de la connaissance et enrichissement de l'information

Si l'on considère la résolution d'un problème tel que la classification d'une base de connaissances, la FIGURE 3.6 illustre comment un raisonneur produit une information complémentaire à la connaissance explicite, au niveau terminologique. C'est de fait la connaissance sur le traitement de la logique qui permet d'expliciter la connaissance terminologique implicite.

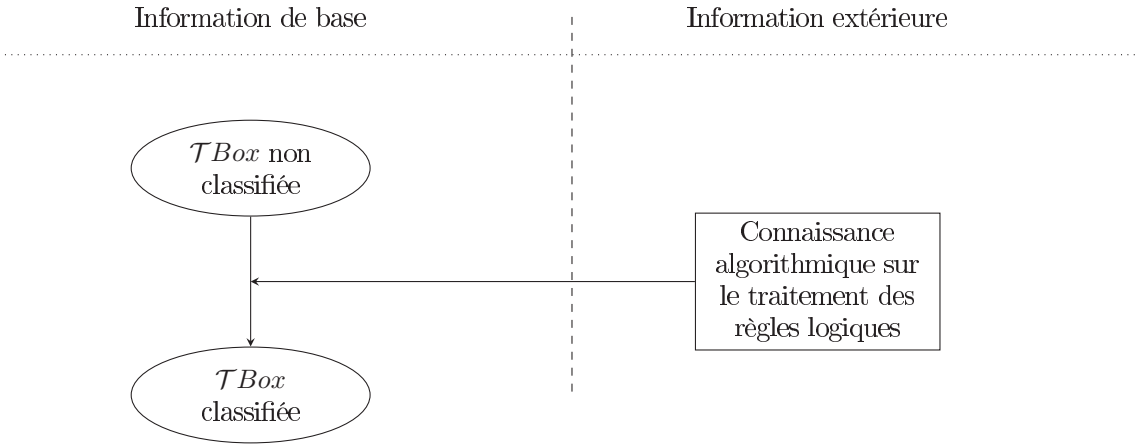


FIGURE 3.6 – Le principe de classification d'une $\mathcal{T}Box$ par un raisonneur.

Pour la réalisation d'une base de connaissances, l'information principale est constituée d'assertions sur des individus. La connaissance terminologique de la base de connaissances est utilisée conjointement à la connaissance sur le traitement des règles logiques pour calculer les concepts les plus spécifiques de chaque individu.

La connaissance terminologique représente une information sur l'information principale : elle lui confère sa structure logique. C'est l'ajout de cette information supplémentaire qui permet d'inférer la réalisation des individus, c'est-à-dire d'enrichir l'information principale (FIGURE 3.7). Nous reviendrons sur ce principe dans notre proposition.

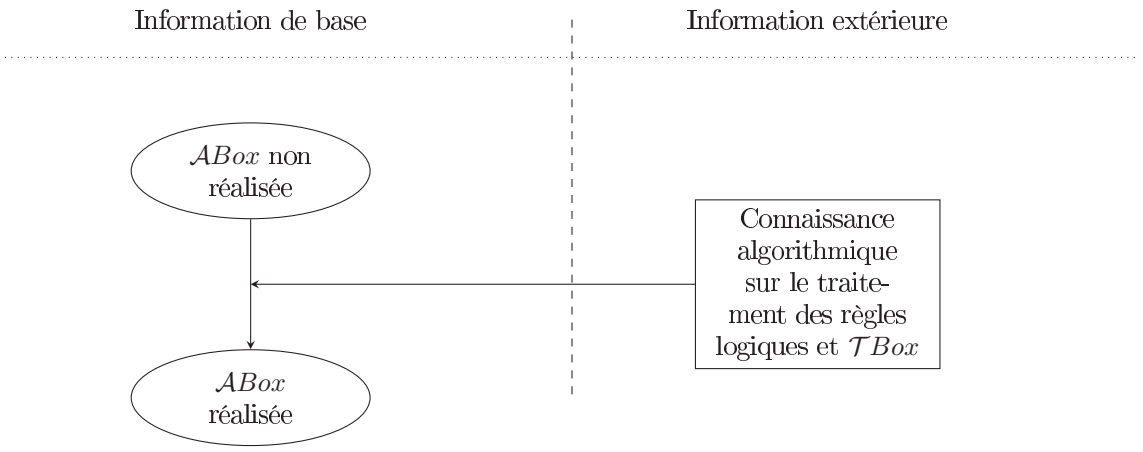


FIGURE 3.7 – Le principe de réalisation d'une $\mathcal{A}Box$ par un raisonneur.

Pour résumer (Apport de la représentation des connaissances)**.** *Les ontologies permettent de représenter un domaine de connaissance et une partie du monde.*

La standardisation des langages proposée par le W3C ouvre la voie au partage *de la connaissance formalisée. Dans ce processus d'échange d'information, on remarque le rôle joué par le métalangage XML dans les langages de sérialisation.*

Le vocabulaire OWL fait référence à des constructeurs basés sur la logique descriptive. En plus de simplement représenter une connaissance explicite, ce langage permet de la partager avec des raisonneurs automatiques. Ceux-ci sont capables d'utiliser des mécanismes d'inférence afin d'expliciter de la connaissance implicite.

Ainsi, la représentation d'une part de connaissance explicite *sous forme logique permet de révéler une connaissance* implicite.

4

Les ontologies en imagerie, télédétection et données spatiales

Dernier chapitre de la partie dédiée à l'exposé préalable, celui-ci est d'abord consacré à l'examen des liens entre image et sémantique dans les paradigmes d'analyse par pixels ou par groupes de pixels. Ceci conduit à se pencher plus en détail sur les différentes approches existantes qui proposent de rapprocher l'image et la connaissance au moyen des ontologies.

Sommaire

4.1 La sémantique des pixels **80**
4.1.1 Sémantique objective et sémantique subjective 80
4.1.2 Sémantique physique et sémantique de télédétection, préclassification de Baraldi et al. (2006) 80
4.2 Analyse d'image orientée objet **82**
4.2.1 Le rôle de la segmentation dans la création des objets 82
4.2.2 Objet géographique et objet de l'image 83
4.2.3 Notion de point de vue 83
4.3 Segmentation et sémantique **85**
4.3.1 Impact de la segmentation sur la sémantique des segments 85
4.3.2 Segmentation guidée par de la connaissance 86
4.4 Les ontologies pour l'annotation d'images **86**
4.5 Formalisation de la connaissance experte **86**
4.5.1 Interprétation sémantique par un algorithme dédié 86
4.5.2 Ontologies et apprentissage 87
4.5.3 Les ontologies en imagerie médicale et autres domaines 87
4.5.4 Conclusion sur les différentes applications 88
4.6 Raisonnement **88**
4.6.1 L'aboutissement de la connaissance formalisée 88
4.6.2 Raisonner sur le contenu d'une image 88
4.7 Structuration des données spatiales **89**
4.8 Réflexions sur la dimension temporelle **90**

Comme en témoignent les recensions effectués par Hichem Bannour (2011) et Dasiopoulou and Kompatsiaris (2010), les ontologies constituent un domaine à la mode dont l'émergence suscite d'importantes perspectives dans le domaine de l'image. Les applications sont nombreuses et si les standards actuels concernent avant tout le web sémantique ils n'excluent pas d'autres domaines d'application. Mieux, le développement historique des raisonneurs a montré que le web sémantique peut même agir comme un levier.

Concernant les images en général et satellitaires en particulier, les ontologies sont perçues comme un moyen de réduire le *fossé sémantique* défini comme :

Définition 64 (Fossé sémantique, d'après Smeulders et al. (2000))**.** Le « manque de concordance entre l'information que l'on peut extraire des données visuelles et l'interprétation que l'on peut en faire dans une situation donnée ».

Par leur capacité à numériser une connaissance terminologique et assertionnelle, les ontologies représentent un sérieux potentiel pour répondre à ce problème en jouant un rôle jonction entre le monde des concepts et le monde de l'image.

Prolongement naturel de l'exploration des disciplines concernées par la thèse, le tour d'horizon bibliographique ne se limite pas aux images satellitaires : même si ce dernier a été privilégié il n'a pas empêché de se référer à des travaux sur les images médicales ou photographiques.

4.1 La sémantique des pixels

4.1.1 Sémantique objective et sémantique subjective

Qui dit « ontologie » dit « sémantique ». Mais de quelle sémantique les images satellitaires qui nous intéressent sont-elles au juste porteuses ?

S'il s'agit de rattacher un pixel au *concept de pixel* ou une image au *concept d'image*, cela ne semble pas vraiment concerner précisément le domaine de la télédétection, ni une application particulière. Cette première sémantique « objective » étroitement liée à une simple *observation* de l'existant, semble assez indépendante d'une façon de voir de chacun. Une image matricielle est composée de pixels... c'est ainsi par définition et il est peu probable que cela soit contesté. Ce genre de conceptualisation peut provenir de définitions, d'*a priori*, de consensus... C'est un niveau de sémantique de base (FIGURE 4.1).

Un autre niveau de sémantique concerne le sens particulier que l'image et son contenu représentent du point de vue de l'observateur, de l'interprète, du télédétecteur. Il s'agit alors d'une sémantique « subjective » qui dépend d'une certaine *interprétation* des entités. Ici, la subjectivité prend le pas sur l'objectivité et les concepts sont davantage soumis à variation en fonction de l'usager qui les définit, des instruments utilisés, de l'objectif poursuivi... bref, d'un certain contexte. Chacun peut avoir une définition différente de ce qu'il considère comme représentant une zone de l'image végétalisée, car l'interprétation de ce terme est personnelle et contextuelle.

La sémantique subjective semble construite à partir d'une sémantique plus objective. En effet, une *zone de l'image représentant de la végétation* (sémantique subjective) est avant tout « une sorte de » *zone de l'image* (sémantique objective). Et l'identité des concepts objectifs de deux entités ne garantit pas l'identité de leurs interprétations plus subjectives (FIGURE 4.2).

La frontière entre la sémantique objective et subjective n'est pas nette et se discute, mais il est important de remarquer que la sémantique d'une entité n'est pas unique *a priori*. Ces considérations rejoignent la distinction que Falomir et al. (2011) proposent entre *conceptualisation de référence* (sémantique de base) et *connaissance contextuelle* (interprétation). Nous y reviendrons par la suite.

4.1.2 Sémantique physique et sémantique de télédétection, préclassification de Baraldi et al. (2006)

Sans chercher pour le moment à délimiter de quelconques objets dans l'image, les données auxquelles elle donne accès se présentent sous la forme d'une suite finie de valeurs. L'ordre des valeurs

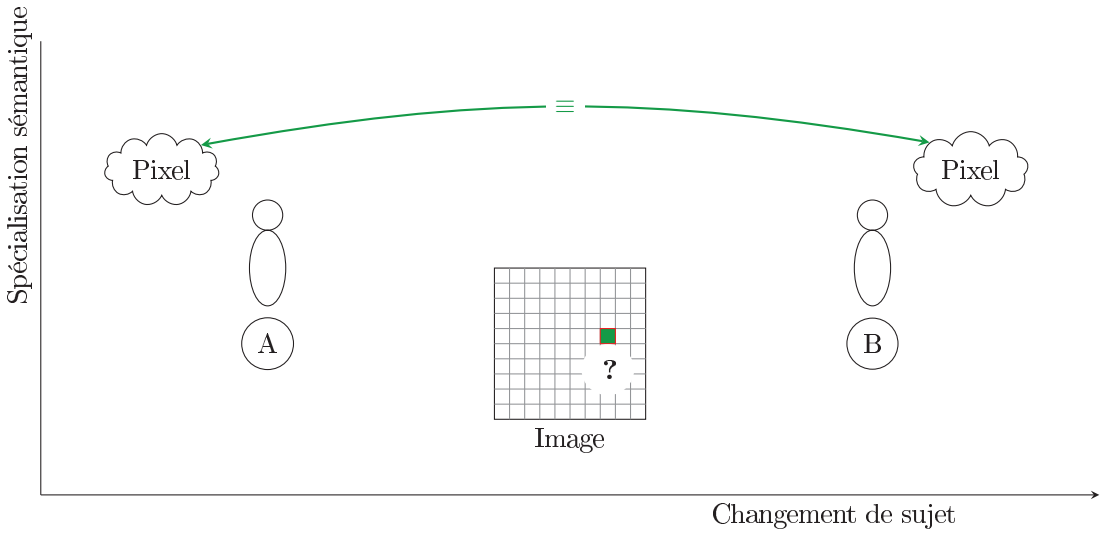


FIGURE 4.1 – Sémantique objective d'une entité. Le fait de rattacher une entité à un concept lui confère une sémantique.

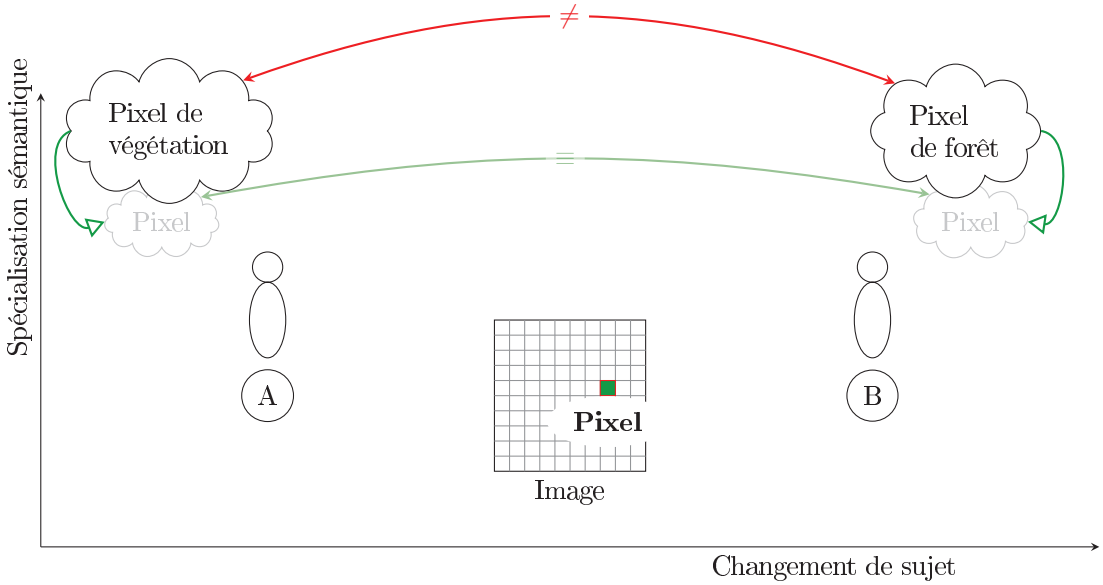


FIGURE 4.2 – Sémantiques subjectives d'une entité. Plus on considère un niveau subjectif de sémantique d'une entité, plus celle-ci est susceptible de séparer les observateurs, même si ceux-ci sont d'accord sur une sémantique plus objective. Ici, deux acteurs s'accordent sur le concept de pixel. Mais ils diffèrent par leur interprétation de ses caractéristiques, ce qui les conduit à lui attribuer des sémantiques subjectives différentes.

renvoie à une position du signal enregistré dans l'espace, ce qui permet ensuite de visualiser l'image. Les valeurs peuvent être scalaires dans le cas d'une image monospectrale ou vectorielles dans le cas d'une image multispectrale (il y a dans ce cas plusieurs valeurs scalaires enregistrées pour une même position dans l'espace). Nous avons vu que chaque position dans l'espace est représentée par un *pixel* (*picture element*, soit littéralement un élément d'image).

Les pixels constituent donc la maille élémentaire de l'information de l'image et il est logique de s'intéresser à leur valeur pour en *déterminer le sens*.

Dans le cas particulier des images satellitaires, la valeur des pixels est un simple compte numérique dénué de signification physique : ce sont de simples valeurs qui ne se réfèrent à aucun concept et qui n'ont donc pas vraiment de sens. Il est toutefois possible de transformer ces valeurs en mesures physiques à condition d'y appliquer une opération de *calibration radiométrique* grâce aux métadonnées habituellement jointes à chaque image. Suite à cette étape, les valeurs des pixels ont véritablement acquis un sens : ce sont des valeurs associées au concept d'une mesure physique particulière (souvent la température ou une réflectance).

Les télédétecteurs manipulent toutefois le plus souvent des concepts plus abstraits que ces grandeurs physiques, comme la couverture végétale ou urbaine, par exemple. Un premier travail académique de préclassification automatique a été proposé par Baraldi et al. (2006) dans le but d'attribuer à chaque pixel une sémantique de télédétection en fonction de sa valeur physique.

Ce travail est basé sur une connaissance experte propre à l'auteur. Celui-ci a construit diverses préclassifications (avec un nombre de classes faible, moyen ou élevé) pour différents satellites (Landsat 5, Landsat 7, Spot 5...). Cela met en évidence que la connaissance experte en télédétection dépend d'une expérience personnelle, des objectifs poursuivis et des caractéristiques du capteur utilisé pour la prise de vue.

En observant l'implémentation d'un algorithme de classification[1], il est aisé de constater que la connaissance utilisée est complexe et que, d'une certaine manière, l'implémentation la fige dans l'algorithme. Cela impose aux utilisateurs de se soumettre aux classes définies par Baraldi et al. (2006), à moins de réimplémenter complètement l'algorithme. On perçoit donc en même temps toute la puissance de cette approche et son manque de souplesse : la connaissance a été définie par une seule personne ; elle n'est pas personnalisable et reste difficilement partageable.

4.2 Analyse d'image orientée objet

Tant qu'une valeur n'est pas associée à un concept, elle n'a pas de véritable sens. Elle en acquiert un, du fait de son étiquetage par un concept. Nous avions auparavant souligné la distinction entre les entités de l'image (qui sont constituées de pixels) et les entités de terrain (qui sont constituées de molécules).

L'augmentation de la résolution spatiale sur les capteurs embarqués à bord des satellites d'observation de la Terre produit une quantité croissante de données (Ruan et al., 2006). Cette évolution semble avoir provoqué l'apparition d'un nouveau paradigme d'analyse des images satellitaires s'appuyant désormais non plus directement sur les briques élémentaires de l'information dans l'image que sont les pixels, mais sur des *groupes* de pixels. On peut voir cette mutation comme le résultat des limites de l'approche pixel (Zhang et al., 2012) ou bien comme l'opportunité de s'intéresser à des objets de dimension de plus en plus réduite qui apparaissent sur les images (Castilla and Hay, 2008).

La pluralité des acceptions des mots pose déjà problème à l'intérieur d'une communauté et la pluridisciplinarité du sujet amplifie encore le phénomène. C'est le cas pour la qualification « orientée-objet ».

4.2.1 Le rôle de la segmentation dans la création des objets

Comme nous l'avons posé au préalable, une *segmentation d'image* est un traitement « partitionnant la totalité de l'image en régions mutuellement disjointes qui présentent une plus grande uniformité vis-à-vis d'elles-mêmes que vis-à-vis des régions adjacentes » (Castilla and Hay, 2008). Le critère de l'uniformité se discute, car on peut toujours découper une image (comme le propose le logiciel *eCognition*) en régions arbitraires... On pourra par contre préciser que chaque région issue de la segmentation (*segments d'image*) est dotée d'un identifiant (Ruan et al., 2006) qui permet de la désigner dans son ensemble comme une sorte d'individu, une entité à part entière.

Parce qu'elle permet l'identification des nouvelles entités que sont les segments, la segmentation est une étape indispensable à toute analyse d'image orientée objet (désignée par le sigle OBIA pour *Object Based Image Analysis*). En dépit du terme « objet », cette méthode n'a rien à voir avec la

[1] Il s'agit en l'occurrence de l'implémentation proposée par le logiciel d'analyse d'images de télédétection *Orfeo Toolbox*, pour les satellites Landsat 5 et Landsat 7, avec un nombre de classes intermédiaire .

programmation orientée objet ni avec les ontologies. Pourtant, dès lors qu'apparaissent des *objets* (au sens très général d'*instances*) et une manière de les regrouper en *classes*, une certaine parenté dans l'abstraction du monde se dessine.

D'autres auteurs, dans Liu et al. (2008), considèrent également la segmentation comme une vectorisation de l'image : les segments peuvent être considérés comme des polygones. Ils insistent aussi sur le fait que les segments sont mutuellement disjoints : les algorithmes de segmentation rattachent un pixel à une région unique (ce qui rejoint pour eux, une notion d'*espace mono-valué*).

Afin de fixer les choses, on peut également définir la *sur-segmentation* par « la situation dans laquelle, d'après l'observateur, le contraste entre quelques segments est insuffisant, ce qui devrait conduire à les fusionner en un unique objet de l'image ». La *sous-segmentation* concerne elle « l'existence de segments qui, d'après l'observateur, manquent de cohérence et devraient être découpés en objets de l'image distincts ».

4.2.2 Objet géographique et objet de l'image

La distinction que nous avons évoquée précédemment entre les entités de l'image et les entités de terrain (Définition 44, page 54) découle principalement de réflexions de Castilla and Hay (2008) sur les conséquences de la création d'objets de l'image par la segmentation.

Castilla and Hay (2008) définissent aussi l'*objet-image* comme « une région discrète d'une image numérique qui présente une cohérence interne et qui diffère de son voisinage ». On peut également dire qu'il s'agit de « segments produits par un *bon algorithme de segmentation* », même si l'algorithme parfait n'existe pas (Durand et al., 2007). Ils remarquent également que les objets de l'image n'ont pas d'existence objective dans l'image, mais qu'ils sont le fruit d'une segmentation particulière. Il s'ensuit que l'OBIA n'*extrait* pas l'information de l'image, mais *produit* l'information à partir de l'image.

Castilla and Hay (2008) distinguent également les objets de l'image des *objets géographiques*, « régions géographiques limitées qui peuvent être identifiées pour une période de temps comme le référent[1] d'un terme géographique ».

Cette distinction entre les objets de terrain et leur représentation sur l'image n'est pas seulement logique, elle est ontologique dans le sens métaphysique du terme[2]. Dans un cas on observe une prairie composée de brins d'herbe ; dans l'autre, on regarde un segment d'image composé de pixels : les natures de ces deux objets n'ont aucun rapport direct. Malgré cette différence fondamentale, le langage courant peut exercer une certaine influence sur notre manière de comprendre les mots et provoquer une sorte de confusion entre une entité physique et ses représentations. En disant dans les deux cas que l'on observe « une prairie », il est facile de rattacher l'objet géographique et l'objet de l'image au même concept. D'autres références soulignent la même distinction, comme Iakovidis et al. (2009) qui dressent ce constat en imagerie médicale.

4.2.3 Notion de point de vue

En un certain sens, cette distinction rejoint en partie ce que Mariño Drews (1993) appelle un *point de vue*[3]. Dans un contexte pluridisciplinaire, cette notion de point de vue acquiert toute son importance. En ce qui nous concerne, elle est susceptible d'intervenir de deux manières.

[1] D'après Castilla and Hay (2008), *en sémiotique, un* référent *est l'objet du « monde réel » que désigne un mot ou un signe, c'est-à-dire une instance du type d'objet. Nous préférons utiliser « référent » plutôt que « instance » parce que le premier met en évidence la distinction entre un objet réel (un objet géographique d'un certain type) et sa représentation (l'objet de l'image classifié comme une instance représentative de ce type d'objet géographique).* La référence directe à la *réalité* peut faire débat dans la communauté scientifique et certains, dans la lignée kantienne, préféreront parler de *phénomènes* issus d'une réalité inaccessible.

[2] Nous voyons au passage que l'acception métaphysique de l'ontologie s'invite naturellement dans les ontologies en informatique du fait que dans les deux cas on s'intéresse aux caractéristiques des choses qui font qu'elles sont ce qu'elles sont.

[3] « La notion de point de vue est absente de la plupart de systèmes de représentation de connaissances à objets. Pourtant les grandes bases de connaissances contiennent en général des objets concernant plusieurs disciplines. Chaque discipline observe et structure les objets selon sa perspective, son point de vue.

[...] Un point de vue détermine une taxinomie de classes et les attributs du concept qui vont être visibles sous ce point de vue. Les différents points de vue peuvent être reliés par une ou plusieurs passerelles... » (Mariño Drews, 1993).

La première concerne la différence de point de vue entre deux observateurs d'un même objet (FIGURE 4.3). Chaque observateur a une connaissance, une expérience, un objectif qui lui sont propres et qui le conduisent à caractériser un objet avec certaines caractéristiques, *a priori* différentes de celles retenues par les autres observateurs. Par exemple, en forêt, tel observateur s'intéressera uniquement à la taille des arbres et tel autre uniquement à leur diamètre ou à leur essence. Ils observent les mêmes objets, mais avec des points de vue différents. Naturellement, un « observateur » s'entend au sens large d'un acteur qui donne une description d'une entité. Cet acteur peut être humain ou informatique. Il peut également observer et décrire successivement la même entité selon des points de vue différents.

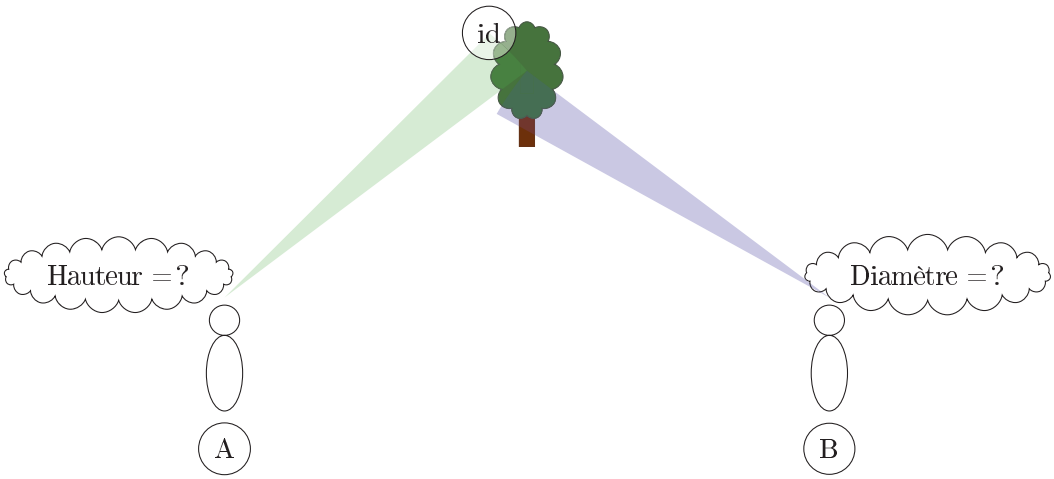


FIGURE 4.3 – Une illustration de la notion de point de vue, d'après Mariño Drews (1993). Celle-ci découle principalement de la différence d'objectifs entre les observateurs (A) et (B) qui les pousse à ne « voir » que les caractéristiques qui les intéressent.

La seconde manière fait intervenir les observateurs d'une entité par des moyens différents (FIGURE 4.4). En télédétection, ce sont le point de vue du terrain (observateur A) et le point de vue de l'image (observateur B). Les deux points de vue se réfèrent à la même entité de terrain (id), mais le point de vue de l'image est indirect et décrit non pas l'entité de terrain elle-même, mais sa représentation dans une image, qui est *une autre* entité (id').

Cette acception de la notion de point de vue, bien que différant de celle de Mariño Drews (1993), la rejoint et l'élargit : il s'agit toujours de deux observations (descriptions) qui se rapportent *in fine* à une même entité (id) et entre lesquelles il est possible d'établir des passerelles.

La FIGURE 4.4 mérite quelques commentaires supplémentaires.

Il y apparaît clairement le fossé sémantique entre les données numériques et le sens qu'elles prennent pour l'observateur (1). On peut également distinguer le *fossé sensoriel* (2).

Définition 65 (Fossé sensoriel, d'après Smeulders et al. (2000); Bissol (2006))**.** Disparité entre un objet physique et sa représentation numérique.

Dans le cadre de la télédétection, le fossé sensoriel est composé de deux parties.

La première est due au fait qu'un capteur (que ce soit l'œil ou un capteur électronique), ne perçoit qu'une partie de l'information en provenance de l'entité observée. On pourrait appeler cette partie « fossé perceptif ». On le trouve entre l'entité observée et le satellite (3a), mais également entre l'entité observée et l'observateur humain sur le terrain (3b).

La deuxième partie est causée par la perte d'information qui a lieu entre le signal reçu par le capteur et l'enregistrement numérique (4). Elle est due à la sensibilité des capteurs, et à la discrétisation du signal. On pourrait la désigner comme « fossé numérique ».

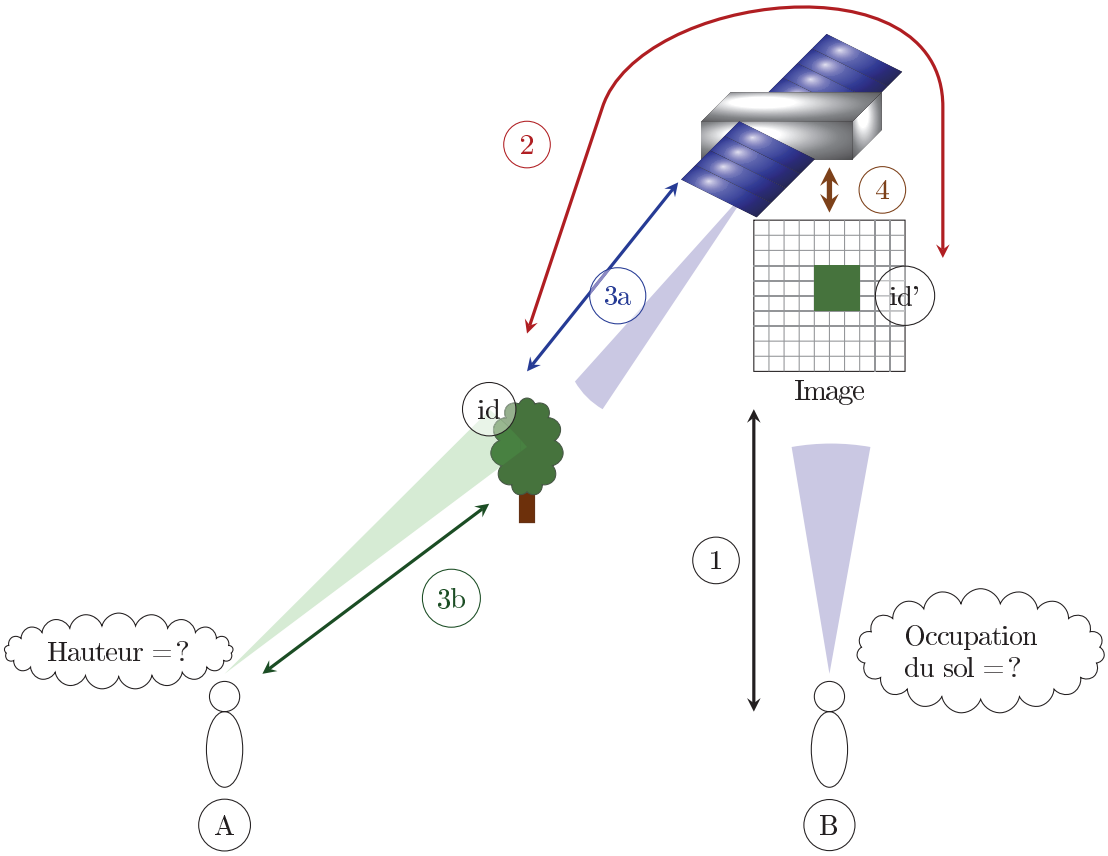


FIGURE 4.4 – Une illustration de l'acception étendue de la notion de point de vue. On y retrouve le *fossé sémantique* entre les données numériques et leur interprétation. On peut également y placer le *fossé sensoriel*. Bien qu'ils s'intéressent à la même entité de terrain (id), les observateurs (A) et (B) décrivent deux entités différentes : respectivement l'entité géographique de terrain (id) et l'entité de l'image (id') qui la représente.

4.3 Segmentation et sémantique

4.3.1 Impact de la segmentation sur la sémantique des segments

L'analyse d'image orientée objet suppose implicitement que la segmentation produit des régions homogènes et significatives, c'est-à-dire représentatives d'objets du monde réel. On cherche en effet la sémantique des entités dont on suppose qu'elles en ont une. La *qualité* de la segmentation exerce une influence directe sur les résultats de l'analyse de l'image orientée objet (Maillot et al., 2004; Schober et al., 2004; Durand et al., 2007; Puissant et al., 2007; Derivaux et al., 2008; Liu et al., 2008; Dasiopoulou and Kompatsiaris, 2010; Falomir et al., 2011). Si les entités dans l'image n'ont pas de cohérence sémantique, alors, soit il sera difficile de les rattacher à un concept, soit le concept auquel elles seront rattachées leur conférera artificiellement une sémantique dont elles ne sont pas réellement porteuses[1]. L'importance de cette étape de segmentation a conduit à des tentatives pour l'évaluer. Zhang et al. (2012) recensent différentes méthodes d'évaluation de la segmentation et en propose lui-même une. D'autres recherches se sont attachées à optimiser le résultat d'un algorithme particulier, en fonction d'une sémantique d'intérêt (Bertrand, 2013).

[1] C'est-à-dire que l'on pourrait se retrouver dans des situations où une zone de l'image qui contiendrait explicitement des surfaces végétales et des surfaces aquatiques pourrait être étiquetée par un concept qui n'aurait rien à voir avec ces sémantiques (par exemple, le concept de zone urbanisée).

4.3.2 Segmentation guidée par de la connaissance

Le développement de nouvelles approches de segmentation constitue à lui seul un domaine de recherche complexe. Son exercice n'est pas trivial pour la plupart des utilisateurs d'algorithmes de segmentation (Simari et al., 2009). On peut trouver de nombreuses façons différentes d'aborder la question. En voici quelques unes dans lesquelles interviennent de diverses manières soit la classification des segments, soit une connaissance experte pour guider le processus, soit le génie ontologique.

Athanasiadis et al. (2007) partent d'une première sursegmentation et procèdent à un pré-étiquetage sémantique. Puis, le résultat est évalué et les régions connexes de même étiquette sont fusionnées. La pertinence du concept affecté à chaque segment est évaluée sur la base d'une distance entre les concepts. Le processus s'arrête lorsque la pertinence atteint un seuil satisfaisant.

Concernant le cas particulier des images de télédétection, d'après Sheeren et al. (2006), le choix des paramètres de segmentation est empirique et le résultat peut être amélioré par l'injection de *données auxiliaires*, en fournissant à l'algorithme une couche thématique de la localisation et de la forme précises des bâtiments.

La segmentation peut donc tirer profit de la connaissance experte dans le découpage des régions. Dans Derivaux et al. (2010), la segmentation est supervisée par une connaissance experte en télédétection qui prend la forme d'exemples. Puis l'algorithme de segmentation peut être appliqué successivement (algorithme génétique) jusqu'à l'obtention d'un résultat satisfaisant.

La connaissance experte nécessaire à la segmentation peut également être contenue dans une ontologie. En imagerie médicale 3D, Hassan et al. (2009) proposent un processus de segmentation hiérarchique. On décompose alors au fur et à mesure les régions découpées et identifiées en régions plus élémentaires jusqu'à aboutir aux feuilles d'une hiérarchie de concepts. Les paramètres de segmentation et les étiquettes à attribuer aux segments sont contenus dans une ontologie. La description des concepts de l'ontologie s'appuie principalement sur des caractéristiques géométriques.

4.4 Les ontologies pour l'annotation d'images

Certains travaux se sont intéressés au génie ontologique pour annoter des images. Dans Hyvönen et al. (2002), l'annotation avec les concepts de l'ontologie est manuelle et permet de naviguer d'une image à l'autre grâce à la proximité sémantique de leurs annotations.

En revanche, l'annotation proposée par Srikanth et al. (2005) est automatique. Les objets d'une image segmentée sont étiquetés par un mot grâce à un algorithme spécifique. Les mots utilisés constituent en fait une hiérarchie de concepts dont la relation d'hyperonymie permet d'annoter les images avec des mots qui n'ont pas été directement trouvés lors de l'analyse. Ainsi, les images de *tigre* ou de *cheval* seront également annotées par le concept d'*animal*.

4.5 Formalisation de la connaissance experte

On le voit, le rapport entre l'image et l'ontologie peut être abordé avec beaucoup de souplesse. C'est toutefois dans un objectif de formalisation de la connaissance experte que les applications semblent les plus nombreuses.

4.5.1 Interprétation sémantique par un algorithme dédié

4.5.1.1 Le projet FoDoMuST

Une large recherche a été menée dans le cadre du projet FoDoMuST[1] (Fouille de données multi-stratégie pour extraire et qualifier la végétation urbaine à partir de bases de données images). Centré sur le milieu urbain, ce projet à conduit à utiliser la connaissance pour guider les algorithmes de fouille et de classification d'images.

Dans Boussaid et al. (2006), la fonction de l'ontologie est encore de modéliser la connaissance experte. À cet effet, cette proposition utilise le standard OWL. Les objets de l'image et leurs caractéristiques sont contenus dans une base de données, mais le choix d'un mécanisme d'appariement

[1] http://acimd.labri.fr/04FICHES/FoDoMuST.htm[accessed : 3-10-2013]

n'est pas arrêté. Durand et al. (2007) qui semblent poursuivre ce travail utilise les caractéristiques spectrales, spatiales, et contextuelles des objets pour les identifier au moyen d'un algorithme spécialement implémenté.

Dans Brisson et al. (2007), l'ontologie contenant la connaissance experte est parcourue par un algorithme d'appariement. Cet algorithme mesure la similarité de chaque « objet géographique » aux différents concepts de l'ontologie. La mesure de similarité est élaborée à partir des caractéristiques de l'image relatives à la forme et à la radiométrie caractéristiques de chaque concept. Le fait qu'un « objet géographique » est caractérisé par des critères de l'image et non pas des critères de terrain, illustre une fois de plus toute l'ambiguïté du langage à propos des entités de terrain et de leurs images.

Toujours dans le domaine des images urbaines, Puissant et al. (2007) utilise les ontologies pour contenir une connaissance experte et insiste tout particulièrement sur le potentiel modulaire des ontologies comme outil de séparation de la connaissance et du traitement[1]. L'interprétation des segments fait toujours appel à un algorithme spécifique qui cette fois-ci tient compte du contexte spatial en plus des caractéristiques spectrales, des indices spectraux et des propriétés de forme.

Les travaux initiés dans le cadre de ce projet sont toujours d'actualité et trouvent leur prolongement, par exemple, dans les publications de Forestier et al. (2012).

4.5.1.2 Autres recherches

Une approche qui semble présenter une certaine parenté sur le principe avait été proposée par Mezaris et al. (2003). Cette proposition voyait les ontologies comme un intermédiaire entre des descripteurs numériques d'une image et des définitions de concepts basées sur des caractéristiques qualitatives. Comme les concepts présents dans l'ontologie, les régions de l'image étaient caractérisées par des descripteurs qualitatifs. L'appariement entre une région et un concept était réalisé en comparant ces descripteurs intermédiaires.

On peut également citer les travaux de Hudelot and Thonnat (2003) qui s'appuient sur des algorithmes spécifiques.

L'ontologie de la connaissance experte utilisée par Derivaux et al. (2008) contient une hiérarchie de concepts urbains pour lesquels des experts ont explicité des contraintes sur les valeurs spectrales. Les caractéristiques des concepts sont prioritairement obtenues par apprentissage sur des exemples, mais dans le cas où ces derniers font défaut, ce sont les intervalles spectraux de l'ontologie qui sont utilisés. Dans tous les cas, c'est un algorithme spécifique qui attribue un concept aux objets de l'image.

4.5.2 Ontologies et apprentissage

La proposition de Maillot and Thonnat (2008) consiste à segmenter des images puis à les annoter manuellement. Le rapprochement entre les mots et les caractéristiques visuelles est obtenu par apprentissage. Une ontologie contient les concepts relatifs aux caractéristiques visuelles (qui semblent qualitatives), ainsi que les relations spatiales de type RCC8.

Les ontologies utilisées par Puissant et al. (2006) pour l'interprétation d'une image urbaine contiennent la connaissance experte. L'appariement des concepts aux objets est réalisé par apprentissage.

4.5.3 Les ontologies en imagerie médicale et autres domaines

Pour Iakovidis et al. (2009), les ontologies permettent de formaliser la connaissance experte sur des images médicales. Cette proposition évoque le potentiel de raisonnement sur la connaissance formalisée, mais ne semble pas en tirer de conséquence opérationnelle. Il s'agit en revanche d'un travail qui distingue explicitement les *objets du monde réel* de leurs représentations.

Dans Hudelot et al. (2008) on trouve une conceptualisation de *relations spatiales floues*. Dans les images médicales dont il est question, les relations spatiales telles que « près de » ou « à droite de » sont considérées comme des concepts.

[1] D'après Mariño Drews (1993), cette séparation de la *connaissance* et du *contrôle* est propre aux systèmes à base de connaissances.

Athanasiadis et al. (2007) décrivent une connaissance de domaine en utilisant RDF. On retrouve là aussi une notion de flou dans les relations liant les concepts entre eux. Pour cela, les auteurs réifient les relations pour ajouter un *degré d'appartenance*.

4.5.4 Conclusion sur les différentes applications

Une spécificité de la connaissance apparaît dans chacun des différents domaines d'application des ontologies. L'importance primordiale de la forme, de la position ou des relations spatiales émerge rapidement dans le domaine de l'imagerie médicale (Hudelot et al., 2008; Hassan et al., 2009). En télédétection, les relations spatiales sont insuffisantes (Puissant et al., 2007; Durand et al., 2007). L'information prise en compte prioritairement est alors la radiométrie[1] des objets (Sheeren et al., 2006).

Tout en utilisant les ontologies principalement comme réceptacle de la connaissance experte en confiant la classification des segments à un mécanisme d'apprentissage, Ruan et al. (2006) introduisent le raisonnement automatique au niveau du requêtage des utilisateurs sur une base d'images analysées.

4.6 Raisonnement

Les ontologies, connues comme un outil de *formalisation de la connaissance* ont souvent été utilisées pour représenter une expertise de domaine. Toutefois, la formalisation de la connaissance n'est pas une transformation neutre : elle a ses avantages et ses inconvénients.

4.6.1 L'aboutissement de la connaissance formalisée

Le principal inconvénient est l'appauvrissement de la connaissance, car les ambiguïtés du langage sont aussi une richesse et peuvent être porteuses d'information et de sens, par exemple par association d'idées.

Le principal avantage est que la connaissance désambiguïsée devient numérisable. À partir d'un certain degré de simplification et moyennant l'utilisation de règles logiques, la numérisation permet d'utiliser des logiciels pour raisonner automatiquement sur la connaissance formalisée, pour en vérifier la consistance ou encore rattacher des instances à des concepts par réalisation. Là, intervient le dilemme entre l'expressivité des logiques de description et leur décidabilité. En effet, la formalisation ne permet pas nécessairement de raisonner automatiquement : un langage « trop » expressif peut être indécidable.

Il n'en reste pas moins que le raisonnement automatique est un potentiel capital de la représentation formelle des connaissances. Contrairement à l'utilisation des ontologies pour représenter une expertise utilisable par des algorithmes dédiés, ce potentiel ne semble pas vraiment émerger en télédétection. Mentionnons toutefois que Forestier et al. (2013) envisagent l'utilisation d'un raisonneur pour des tâches bien déterminées telles que la recherche de régions de l'image qui auraient été mal interprétées par un algorithme dédié.

Quelle est la raison de cette réserve ? Peut-être la complexité des logiques de description au regard de la quantité de données spécifique à la télédétection ? Peut-être tout simplement la phase de développement de laquelle le génie ontologique n'est pas encore complètement sorti ?

Pourtant, en dehors des spécificités de la télédétection, il existe des travaux qui ont exploré l'idée de raisonner sur le contenu de l'image. Ils méritent d'être mentionnés.

4.6.2 Raisonner sur le contenu d'une image

Neumann and Möller (2008) utilisent les logiques de description (semble-t-il sans s'appuyer sur un standard ontologique) pour décrire une scène et raisonner sur sa représentation.

Schober et al. (2004) s'appuient sur une ontologie de domaine basée sur des standards[2]. Un apprentissage permet de caractériser les concepts à l'aide de caractéristiques qualitatives de l'image

[1] La radiométrie désigne la mesure du rayonnement électromagnétique, en particulier dans le domaine de la lumière visible (photographies de paysages) mais également dans d'autres longueurs d'onde (images satellitaires).

[2] Le standard DAML+OIL, quelque peu désuet aujourd'hui, peut être considéré comme un ancêtre d'OWL.

et de relations spatiales. La classification est réalisée par un raisonneur sur la base d'une description qualitative des objets et de l'ontologie de domaine.

C'est sans doute Falomir et al. (2011) qui poussent le plus loin le potentiel de raisonnement des logiques de description combinées aux standards du génie ontologique. Ils s'appuient sur une *conceptualisation de référence* (ou ontologie cadre) pour représenter la connaissance de base valable quel que soit le contexte d'application et sur une *connaissance contextualisée* (ou ontologie de domaine) pour expliciter une connaissance relative à une application particulière. La connaissance contextuelle utilise la conceptualisation de référence, mais l'approche n'en reste pas moins modulaire. À un degré moindre, cette modularité apparaît aussi dans Koletsis (2009), voire dans Ruan et al. (2006). Dans l'exemple proposé, le contexte consiste à définir les concepts de *mur*, *sol*, *porte de bureau* et *extincteur* que l'on cherche à identifier sur des images de couloirs. L'image est d'abord segmentée, et chaque objet-image caractérisé par des propriétés issues de l'image (couleur, forme...). Un *modèle qualitatif* permet de transformer la description quantitative des objets en termes qualitatifs. La description qualitative est ensuite représentée sous forme ontologique en peuplant la conceptualisation de référence : ce sont les *faits de l'image*. Cette représentation en *trois couches* (conceptualisation de référence, connaissance contextuelle et faits de l'image) alimente un raisonneur qui rattache les individus de l'image à des concepts dépendant du contexte. Cette proposition exploite l'expressivité de la logique de description $\mathcal{SHOIQ}$ à travers le standard OWL.

4.7 Structuration des données spatiales

Enfin, les ontologies sont un support de connaissance qui a fait son entrée dans la communauté de l'information géographique. Par exemple, *LinkedGeoData* et *GeoLinkedData* sont deux projets (respectivement allemand et espagnol) de représentation des données géographiques en RDF. Il s'agit là encore d'une conséquence du Web sémantique dont le développement est encadré par des organismes de standardisation et de normalisation : l'ISO[1], l'OGC[2] et le W3C (Koubarakis et al., 2012).

La description des données spatiales fait appel à divers modèles concernant les géométries et les relations spatiales (Koubarakis et al., 2012). Dans le contexte des relations spatiales, Clementini (2009) utilise le terme de *relation* dans un sens compatible avec la Définition 53 et voit dans la géométrie des objets, une *relation unaire*. Selon ses travaux, les relations spatiales peuvent être classées d'une part selon leur nature *topologique*, *projective* ou *métrique* et d'autre part en fonction du nombre d'entités concernées par la relation : *unaire*, *binaire*, *ternaire* et *n-aire*.

Koubarakis et al. (2012) recensent de nombreux vocabulaires et ontologies réutilisables : l'ontologie géopolitique de la FAO[3], l'ontologie GML[4] etc. Leur article examine plus particulièrement la représentation des relations spatiales (ou plus exactement de l'information spatio-temporelle) avec RDF. Il examine des propositions qui se ramènent à l'intégration de deux types d'information. D'abord, l'aspect spatial est représenté par la définition de types de géométrie. Mais c'est surtout l'aspect temporel qui étend la structure des triplets en y ajoutant un instant ou un intervalle de temps de validité.

D'autres approches sont signalées dans cette recension, comme les travaux de Kolas and Self (2007) qui intègrent l'information spatiale sans extension des standards, en utilisant simplement un vocabulaire adéquat à la représentation des relations spatiales (topologiques). Ceci démontre que la représentation des relations spatiales dans une base de connaissances n'implique pas forcément l'élaboration d'un modèle d'ensemble spécifique, mais simplement la définition d'une terminologie. Comme pour tout autre domaine de connaissance, ce vocabulaire peut enrichir une conceptualisation existante.

Quoi qu'il en soit, la représentation des relations spatiales s'appuie tout de même sur une modélisation préalable. Le travail de Clementini (2009) en ce domaine a été cité plus haut, mais Egenhofer (1989) et Randell et al. (1992) (RCC8[5]) avaient également proposé des représentations

[1] *The International Organisation for Standardization* : http://www.iso.org.
[2] *The Open Geospatial Consortium* : http://www.opengeospatial.org/
[3] *Food and Agriculture Organisation of the United Nations* : http://www.fao.org/countryprofiles/geoinfo.asp?lang=en.
[4] *Geography Markup Language* : http://loki.cae.drexel.edu/~wbs/ontology/2004/09/ogc-gml.owl.
[5] *Region Connection Calculus*... ou bien d'après les initiales des auteurs.

formelles des relations topologiques spatiales binaires.

En ce qui concerne plus spécifiquement le potentiel d'OWL pour la représentation de l'information spatiale, plusieurs pistes existent (Hogenboom et al., 2010). Katz and Grau (2005) proposent d'utiliser les logiques de description pour décrire les relations entre des régions vues comme des ensembles de points. Dans le cadre des logiques de description, les points sont assimilés à des individus et les régions à des concepts. D'autres travaux vont jusqu'à intégrer des relations spatiales topologiques (RCC8) et projectives (CSD9[1]) dans un prototype de raisonneur (Batsakis and Petrakis, 2010; Christodoulou et al., 2012).

4.8 Réflexions sur la dimension temporelle

Comme cela affleure dans les travaux sur les questions temporelles abordées par Koubarakis et al. (2012), les questions spatiales sont liées de diverses manières aux questions temporelles. Randell et al. (1992) évoquent pour leur part le lien entre les relations topologiques spatiales et la proposition développée à la suite de Allen (1983) dans le domaine temporel.

Casati et al. (1998) posent de nombreuses questions relatives à la représentation ontologique des *changements* qui pourraient affecter les entités géographiques dans le temps. À première vue, la question du changement laisse une impression paradoxale (Klein, 2007) : le *changement* d'une entité nécessite la *persistance* de son identité dans le temps.

- Soit l'entité change au sens ontologique, c'est-à-dire que son identité est modifiée. Dans ce cas, l'entité n'a pas changé, mais a été remplacée par une autre entité qui lui a succédé dans le temps. C'est la vision attribuée à Héraclite.
- Soit l'entité ne change pas au sens ontologique, c'est-à-dire que son identité n'est pas modifiée. Dans ce cas, les propriétés de l'entité en question peuvent être modifiées. On rattache ce courant d'opinion à Parménide.

C'est ainsi que Castilla and Hay (2008) relèvent qu'une entité géographique peut évoluer au cours du temps (au sens de la modification de ses propriétés et de la constance de son identité) ou bien être remplacée par une autre entité.

Les ontologies en informatique, par leur large usage de définitions intentionnelles, se basent sur des propriétés pour définir les types (concept, classe) des entités ainsi caractérisés de manière statique. D'autre part, une entité est typée de manière définitive[2].

Une ontologie est donc une structure essentiellement statique dans laquelle la représentation d'éventuelles modification des propriétés et des types des entités au cours du temps n'est donc pas un problème trivial. La Tbox qui contient des vérités générales, comme la Abox qui contient des vérités circonstancielles sont toutes deux des représentations fondamentalement statiques.

[1] *Cone-Shaped Directional logic formalism.*

[2] Ceci est également vrai pour une définition extensionnelle.

Pour résumer. *Dans ce chapitre, nous avons distingué différents niveaux de sémantique d'une même entité, plus ou moins objectifs, plus ou moins liés à la subjectivité de l'observateur. La limite entre objectivité et subjectivité se situe là où commence l'*interprétation *et cette limite peut faire l'objet de discussions. Pourtant, elle permet de distinguer d'une part des éléments conceptuels de base et et d'autre part des éléments conceptuels définis par l'interprète, à partir des éléments conceptuels de base.*

Cela nous a amené à évoquer les travaux de Falomir et al. (2011) et de Baraldi et al. (2006), qui se penchent tous deux, quoique de manière totalement différente, sur le passage d'une sémantique de base à une sémantique interprétée.

Dans son principe, la proposition de Baraldi et al. (2006) affecte une sémantique aux pixels de l'image. Il s'agit-là d'un des deux grands paradigmes d'analyse, l'autre étant l'analyse orientée-objet, c'est-à-dire relative à des groupes de pixels cohérents plutôt qu'à des pixels isolés.

Nous sommes revenus sur la segmentation, en précisant son rôle capital dans ce deuxième paradigme d'analyse. Puis nous avons situé différentes pertes d'information en télédétection sur un schéma élargissant la notion de point de vue proposée par Mariño Drews (1993).

Ces considérations furent suivies de l'examen du rôle de la connaissance à différents stades de l'exploitation des images (segmentation, classification, annotation) dans différents domaines (télédétection, imagerie médicale) et de diverses manières (algorithme dédié, raisonneur).

C'est justement l'utilisation du raisonnement automatique qui semble pousser le plus loin le potentiel offert par la formalisation d'une connaissance conceptuelle.

Deuxième partie

Proposition d'approche

5

Principe général de l'approche

Étant donné que la thèse s'articule autour du rôle que les ontologies pourraient jouer dans l'interprétation sémantique des images satellitaires, il semble intéressant de s'appuyer sur la vision proposée par Falomir et al. (2011). Celle-ci exploite au mieux le potentiel qui découle de la formalisation même de la connaissance : le raisonnement automatique basé sur les logiques de description. Après un développement empirique (Andrés, 2012), cela permet de dégager une première structuration de l'approche composée de trois descriptions (Andrés et al., 2012c,b) : la *conceptualisation de référence*, la *connaissance contextuelle* et les *faits de l'image*, pour reprendre le vocabulaire posé par Falomir et al. (2011).

Chacune des trois descriptions est élaborée d'une manière particulière. Ce chapitre détaille leur construction avant de proposer le schéma de réflexion du principe de l'approche selon lequel elles sont susceptibles de contenir implicitement *une interprétation* d'*une image*, déterminée par *une expertise*.

Sommaire

5.1 Conceptualisation de référence **96**
5.1.1 Ontologies de la conceptualisation de référence 96
5.1.2 Concurrence des conceptualisations de référence 100
5.1.3 Objet de la conceptualisation de référence 102
5.2 Connaissance contextuelle **103**
5.3 Peuplement par la description de l'image **105**
5.4 Interprétation sémantique de l'image **106**
5.4.1 Opération d'union sur les bases de connaissance 106
5.4.2 Analogie avec la règle de trois et le produit en croix 106

5.1 Conceptualisation de référence

Définition 66 (Conceptualisation de référence, adaptée d'après Falomir et al. (2011))**.** Une conceptualisation de référence est une connaissance à propos de concepts élémentaires, qui peut être considérée comme valide par une part plus ou moins large de la communauté d'un domaine, indépendamment de l'application qui peut en être faite.

Définition 67 (Contexte de validité (d'une conceptualisation ou d'une connaissance))**.** Le contexte de validité désigne le domaine d'application dans lequel une conceptualisation peut-être considérée comme correcte, que ce soit par une large communauté ou un nombre réduit d'experts.

Une conceptualisation de référence est valide dans un domaine relativement large et bien déterminé. Par exemple, dans le domaine de l'image satellitaire toute image devrait idéalement pouvoir être décrite suivant une unique conceptualisation de référence.

Les concepts de base décrivant le domaine de l'imagerie satellitaire sont contenus dans la conceptualisation de référence. Comme référence, ils sont généralement *relativement* partagés... du moins par les acteurs censés les utiliser dans les deux autres descriptions qui en dérivent :

- dans la connaissance contextuelle, par *spécialisation*[1] ;
- dans les faits de l'image, par *instanciation*.

Malgré tout, une conceptualisation de référence, comme modélisation, reste une *façon de voir les choses*. En dépit de sa prétention à une certaine objectivité, elle dépend tout de même d'un point de vue. Dans l'absolu, chaque acteur du système peut avoir sa propre conceptualisation de référence, ses propres concepts élémentaires.

On distinguera donc la conceptualisation de référence *propre à chaque acteur*, de la conceptualisation de référence *globale du système*. Ainsi, plusieurs conceptualisations de référence peuvent coexister. Le passage d'une conceptualisation de référence à l'autre sera alors d'autant plus aisé qu'elle s'appuieront sur une même ontologie de haut niveau comme trame commune de leurs visions du monde.

La conceptualisation de référence du système est constituée des conceptualisations de référence particulières aux acteurs du système, ainsi que des correspondances établies entre elles[2].

La conceptualisation de référence utilisée (et pour partie construite) dans le cadre de la thèse, se présente comme une ontologie, ou plus précisément un ensemble d'ontologies, que nous allons détailler.

5.1.1 Ontologies de la conceptualisation de référence

Afin de prendre en compte la possibilité de différentes conceptualisations de référence, nous avons distingué deux acteurs : la brique logicielle et l'utilisateur humain (expert en télédétection). Les extraits d'ontologies dans cette partie (FIGURE 5.1 et FIGURE 5.2) concernent la brique logicielle, c'est-à-dire l'outil d'analyse et de traitement d'image. La modélisation de la structure et des traitements de l'image satellitaire découle principalement des classes constituant l'interface de programmation de cet outil.

5.1.1.1 Structures de l'image satellitaire

La première partie de la conceptualisation de référence concerne la description structurelle des données relatives au domaine de l'image satellitaire. Les concepts structurels de l'image satellitaire

[1] Ou par équivalence, c'est-à-dire double spécialisation.
[2] Cet aspect concerne les questions d'alignements d'ontologies.

sont de deux sortes. Ce sont d'une part les concepts qui se réfèrent à une emprise spatiale dans l'image et d'autre part les concepts relatifs aux valeurs associées à ces emprises.

Les concepts d'image, d'image multispectrale, de segment d'image, de pixel etc. appartiennent à la première catégorie. Ils font tous référence à l'espace qui accueille les valeurs des données, ou plus précisément l'espace référencé par ces valeurs.

Les valeurs elles-mêmes sont représentées de diverses façons. Celles associées aux pixels sont souvent littérales. Il en va autrement des segments dont les valeurs sont plus complexes et représentées par un concept dédié à la description des segments : *SegmentFeature*. Il s'agit d'une abstraction de conteneur de valeurs littérales relatives :

- soit à des indices statistiques calculés par rapport à la valeur des pixels d'un segment d'image pour une caractéristique donnée (référencés par une instance de *StatisticsSegmentFeature*) ;
- soit à des indices de forme décrivant les contours d'un segment produit par une segmentation (référencés par une instance de *ShapeSegmentFeature*).

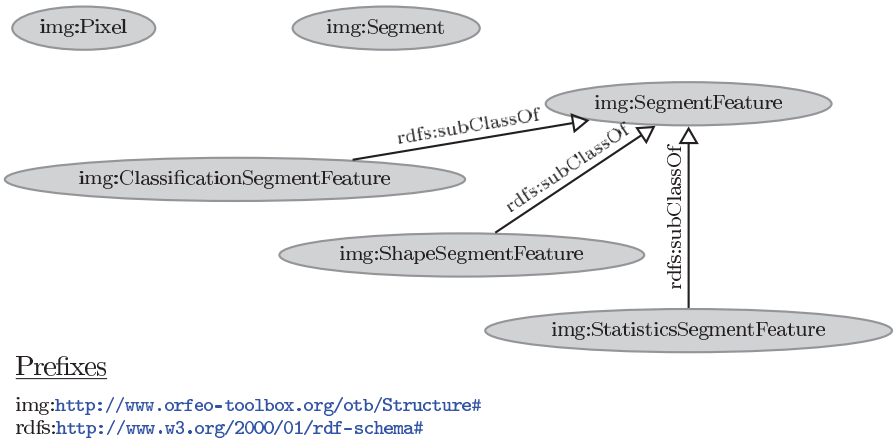


FIGURE 5.1 – Un extrait de la conceptualisation de référence : les principaux concepts relatifs à la structure de l'image utilisés par le logiciel d'analyse d'images de télédétection dans le cadre de la thèse.

5.1.1.2 Traitements de l'image satellitaire

La conceptualisation des valeurs et de leur emprise spatiale ne suffit pas à la description du domaine de l'image satellitaire. Il faut pour cela associer les valeurs à une *caractéristique*, de manière à ce qu'elles prennent un sens pour l'observateur. En télédétection, ces caractéristiques découlent la plupart du temps de traitement.

Si l'on regarde la valeur d'un pixel dans une bande spectrale donnée, on peut considérer que l'extraction de bande qui a produit l'image caractérise la valeur du pixel en l'associant de fait à une longueur d'onde bien déterminée (celle du canal considéré). De même pour un indice spectral, la valeur produite n'est significative que si elle est rapprochée du traitement qui l'a générée. Par exemple, la valeur 0.2 d'un pixel n'a pas de sens si l'on ne précise pas qu'elle a été produite par calcul du NDVI. On peut raisonner de la même manière pour une valeur de texture etc.

Pour représenter cette acception de la sémantique des valeurs, il est nécessaire de poser les concepts relatifs aux différents traitements applicables aux images satellitaires. Initialement posés dans la même ontologie que les structures, il est apparu intéressant de les séparer, de manière à favoriser la modularité à moyen terme. En effet, l'objectif de la thèse n'est pas de proposer une description exhaustive de la structure ni des traitements des images de télédétection. Il est donc important d'anticiper les possibilités de remplacement des ontologies proposées par des conceptualisations plus poussées ou simplement plus complètes. Mais cette séparation logique a ses limites du fait de l'imbrication étroite de la structure et des traitements de l'image.

Les concepts structurels de l'image et les traitements ont vocation à collaborer via des propriétés. Par exemple, une image de segmentation est une image (un concept structurel) qui se réfère assez intuitivement au traitement de segmentation qui l'a produit. Si l'on considère que les traitements ont vocation à définir des structures d'images spécialisées comme l'image de segmentation, il semble logique d'importer l'ontologie des traitements dans l'ontologie des structures (puisque c'est la seconde qui utilise la première). Inversement, on peut imaginer que l'ontologie de l'image serve à définir des traitements spécialisés, en fonction, par exemple, des types de données en entrée et en sortie. Dans ce cas, il serait préférable d'importer l'ontologie des structures dans l'ontologie des traitements.

On le voit, cette question est moins simple qu'il ne semble de prime abord. Il est logique de séparer la représentation des structures de celle des traitements, mais les deux sont intimement liées. Pour le moment, la problématique de la thèse étant davantage centrée sur les données que sur les traitements, c'est la première solution qui a été adoptée. Mais peut-être est-il tout simplement préférable de laisser ces deux conceptualisations indépendantes l'une de l'autre et de définir dans d'autres ontologies les axiomes qui les utiliseraient de concert.

Précisons enfin que les traitements *concrets* sont représentés non pas par des concepts, mais par des individus. On considère ainsi que le calcul du NDVI sur deux images satellitaires fait appel au *même* traitement[1]. Cette vision des choses part du principe qu'une instance de traitement est liée à sa définition algorithmique (implémentée ou non) accompagnée de paramètres bien définis. Quel que soit le nombre d'applications d'un traitement donné, il concerne toujours la même instance.

Toutefois, tous les traitements n'ont pas forcément vocation à être représentés *a priori* dans l'ontologie des traitements. Un certain nombre d'indices peuvent être définis par défaut car ils sont parfaitement déterminés, comme le NDVI. D'autres traitements seront définis à la volée par la valeur de leurs paramètres. Par exemple, par opposition au NDVI qui est parfaitement défini sans paramètre, les indices de textures d'Haralick sont de nature différente : seule la formule est parfaitement déterminée tandis que les autres paramètres sont laissés à l'appréciation de l'utilisateur. De très nombreuses instances de traitements peuvent donc être définies, se référant à une même formule. Il en va de même pour d'autres traitements paramétrés telles que les opérations de segmentation traditionnelles de partage des eaux ou de déplacement de la moyenne.

5.1.1.3 Relations spatiales

Enfin, d'autres conceptualisations cadres peuvent intégrer la conceptualisation de référence, qui dépassent le domaine des images satellitaires. C'est le cas des relations spatiales que nous utiliserons sous la forme de relations entre entités de l'image (par exemple, l'adjacence de deux segments d'image)... mais qu'on peut également imaginer utiliser pour relier des entités de terrain (par exemple, l'adjacence d'une route et d'un champ).

Nous l'avons vu dans l'état de l'art, les travaux sur les relations spatiales ont connu un développement important. En particulier, les standards s'y sont intéressés. L'environnement technique de la thèse a conduit à considérer les relations spatiales RCC8 avec une attention particulière. Nous verrons que la bibliothèque *Orfeo Toolbox*, utilisée dans la mise en œuvre pour le traitement des images satellitaires, implémente le calcul de ces relations. C'est pour cette raison qu'elles ont été explorées de préférence aux autres. Il n'en est pas moins vrai que les considérations générales sur les relations spatiales topologiques dépassent le cas particulier de RCC8.

Or, les relations RCC8 sont représentées dans l'ontologie GeoSPARQL sous la forme de propriétés. GeoSPARQL est un standard de l'OGC reconnu[2]. Ce choix de représentation a donc été adopté, mais dans un souci de simplicité, les huit relations spatiales ont été implémentées dans une nouvelle petite ontologie. Comme les structures sont identiques, cette ontologie provisoire peut être remplacée à tout moment par l'ontologie GeoSPARQL[3].

[1] Identité au sens de l'instance. On pourrait dire qu'il y a un seul traitement NDVI qui peut être appliqué plusieurs fois.

[2] GeoSPARQL n'implémente pas seulement le modèle de relations spatiales topologiques RCC8, mais également celles proposées dans le standard *Simple Features* (Herring, 2011), ainsi que le modèle proposé par Egenhofer (1989).

[3] Cette implémentation d'une petite ontologie évite de multiplier les importations de nombreuses et diverses ontologies utilisées dans GeoSPARQL, telles que le *Dublin Core*, *Skos* ou GML. D'un point de vue méthodologique, ces terminologies supplémentaires n'apporteraient rien à l'utilisation des relations spatiales topologiques, mais poseraient des difficultés d'ordre technique, en particulier concernant la lecture des fichiers avec un parseur encore en

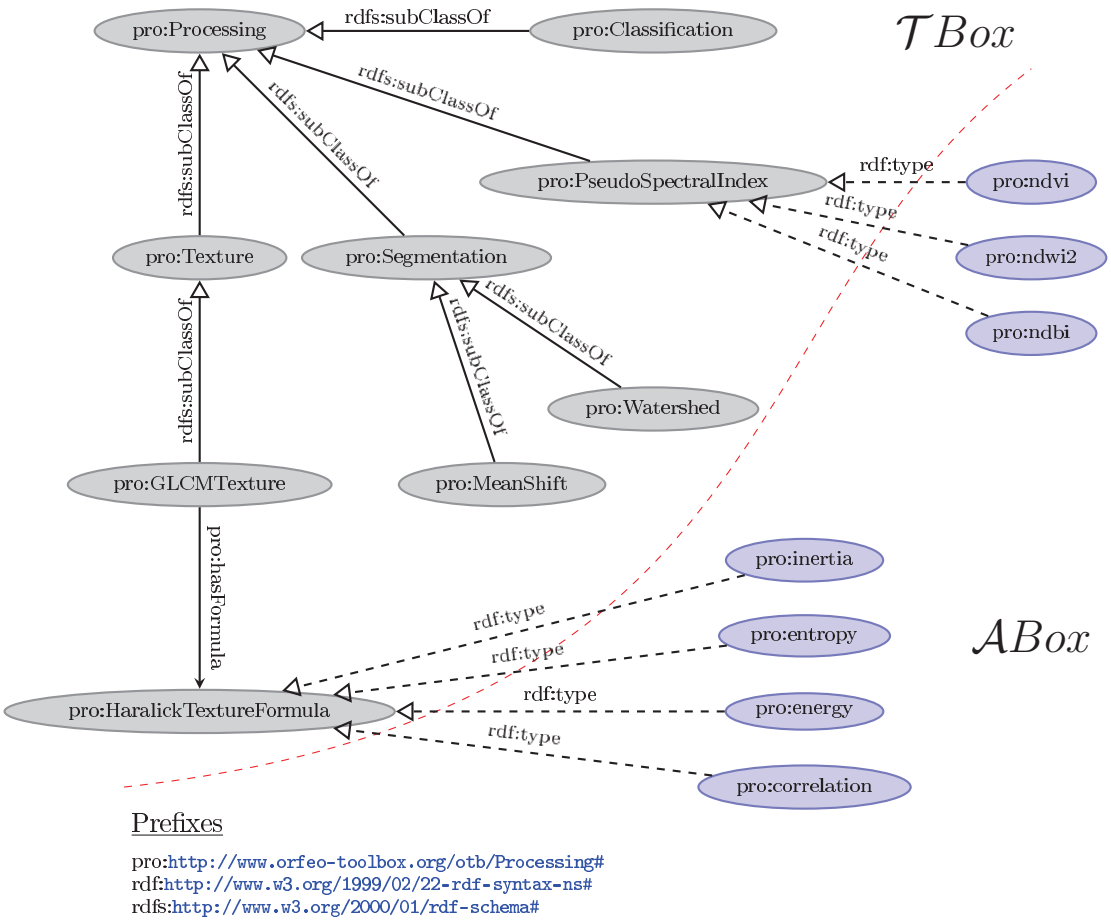


FIGURE 5.2 – Un extrait de la conceptualisation de référence : les principaux concepts relatifs aux traitements de l'image utilisés par le logiciel d'analyse d'images (*Orfeo Toolbox*) de télédétection dans le cadre de la thèse.

5.1.1.4 Partage de la connaissance : OBOE

Il est également intéressant de rattacher la conceptualisation de référence à une ontologie de plus haut niveau.

Ceci a deux principaux avantages. Premièrement, on peut ainsi envisager sereinement le lien avec d'autres conceptualisations de référence basées sur la même ontologie de haut niveau. Dans le cas précis qui nous concerne, on pense en particulier à la liaison entre le monde de l'image et les observations de terrain. Deuxièmement, on bénéficie par ce biais, d'une réflexion globale et réfléchie, d'une certaine *façon de voir le monde* cohérente et partagée par les concepteurs de l'ontologie de haut-niveau.

Le choix d'une conceptualisation de haut-niveau s'est fait en concertation avec un groupe de travail chargé de la modélisation ontologique du terrain basée sur les spécifications LCCS (*Land Cover Classification System*) et GHC (*General Habitat Categories*). Plusieurs candidats ont été examinés, tels l'ontologie SWEET (*Semantic Web for Earth and Environmental Terminology*). C'est finalement l'ontologie OBOE (*Extensible Observation Ontology*) pour des raisons de familiarité technique permettant une prise en main plus rapide de cette conceptualisation. Ce choix qui concerne principalement des questions de partage et de communication entre équipes n'affecte en rien le principe de l'approche.

Le principe d'OBOE est centré autour de l'*observation* d'*entités* (FIGURE 5.3). Chaque observation d'une entité donne lieu à des *mesures* associant chacune une *valeur* à une *caractéristique*,

développement.

étant donné qu'une valeur est considérée comme une *entité*. La combinaison d'une valeur avec une caractéristique correspond tout-à-fait à la façon avec laquelle on peut voir la signification des traitements par rapport aux données des images satellitaires.

À noter que l'extension spatiale d'OBOE, comme ontologie de haut-niveau, intègre une conceptualisation spatiale. Elle propose en particulier un concept de *donnée spatiale* qui spécialise le concept d'*entité*.

Cette extension contient également des concepts permettant de réifier les relations spatiales entre des entités. Pourtant, ce dernier aspect de la conceptualisation n'a pas été utilisé. Une première raison à ce choix en est que la réification des relations spatiales génère autant d'individus qu'il y a de relations, ce qui ne peut que nuire aux performances. Mais la cause déterminante est plus fondamentale. La principale relation spatiale considérée dans l'image est l'adjacence, qui est symétrique. Or, l'ontologie OBOE n'explicite pas la manière avec laquelle il convient d'exprimer la relation et les différents essais de représentation ont tous mené à des incohérences voire à des inconsistances de la base de connaissance. Ces possibilités seront présentées en Annexe B.

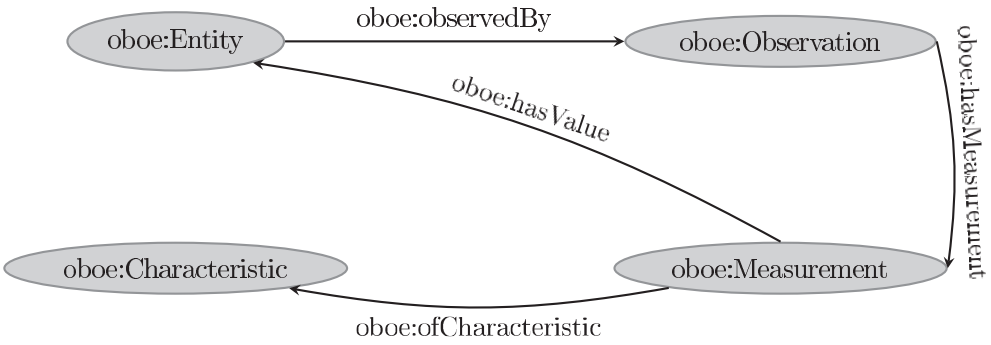


Figure 5.3 – Un extrait de l'ontologie OBOE.

5.1.2 Concurrence des conceptualisations de référence

5.1.2.1 Concurrence des visions de la brique logicielle et de l'utilisateur humain

La partie précédente présentait la modélisation ontologique du côté de la brique logicielle. La vision des choses de l'utilisateur pourrait être le fruit d'une réflexion indépendante, mais dans le cadre de la thèse, elle a été élaborée parallèlement à la conceptualisation incluse dans la brique logicielle et grâce à elle. La conceptualisation de référence de l'acteur « brique logicielle » et celle de l'acteur « utilisateur humain » sont donc forcément très proches. Dans les ontologies, par exemple, le concept de segment du point de vue de la brique logicielle et le concept de segment du point de vue de l'utilisateur humain se distinguent par leurs espaces de noms. De plus, les deux visions s'appuient sur OBOE.

En tant que niveau d'abstraction le plus élevé, OBOE détermine fortement la vision des choses des ontologies d'un niveau d'abstraction moins élevé. Elle contraint en particulier la manière de voir la structure de l'image satellitaire et les traitements de télédétection, car cette connaissance de domaine doit rester compatible et former avec OBOE un tout cohérent, comme l'illustre l'extrait de la conceptualisation de référence de l'utilisateur présenté en Figure 5.4.

Dans le cadre de la distinction formelle des conceptualisations de la brique logicielle et de l'utilisateur humain, la correspondance entre les deux familles d'ontologies est définie par des équivalences de concepts et de rôles élémentaires, explicitant par exemple que les concepts de *Segment* pour le logiciel et de *Segment* pour l'utilisateur sont équivalents.

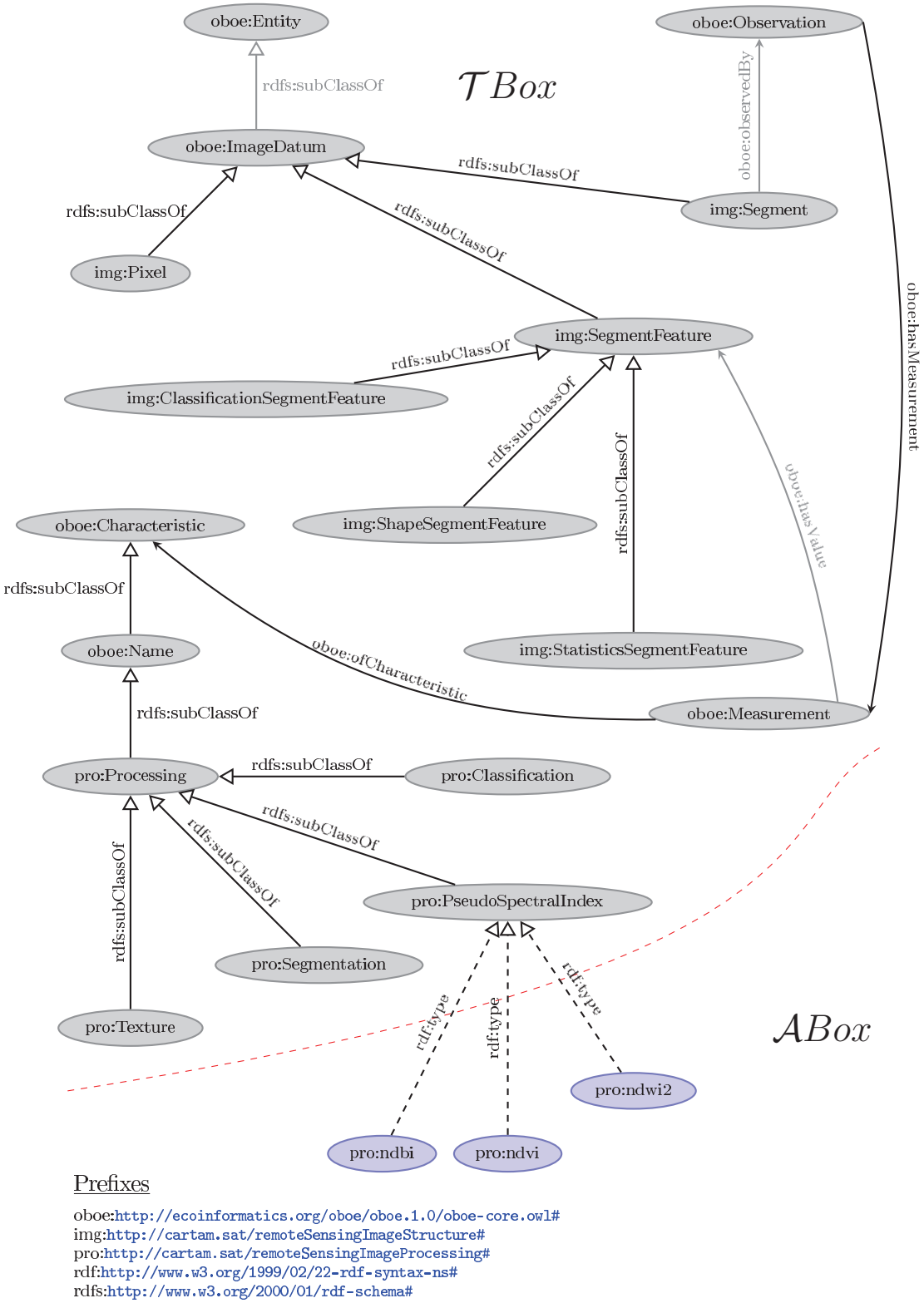


FIGURE 5.4 – Un extrait de l'ontologie de conceptualisation de référence de l'expert. Par opposition, la FIGURE 5.1 et la FIGURE 5.2 étaient relatives à la conception de l'image du point de vue de la brique logicielle. Cette différence apparaît dans les différences d'espaces de noms utilisés dans ces deux dernières figures, et la représentation ci-dessus.

5.1.2.2 Concurrence d'autres ontologies de haut niveau

Le choix de l'utilisation d'OBOE a ses avantages et ses inconvénients et d'autres conceptualisations de plus haut-niveau d'abstraction ont été envisagées, ainsi que la possibilité de ne s'appuyer sur aucune ontologie plus abstraite que la conceptualisation de l'image satellitaire.

Les avantages qui découlent d'un niveau d'abstraction plus élevé sont classiques en génie ontologique : les passerelles avec d'autres applications en télédétection pourraient s'en trouver facilitées. Or, il se trouve que d'autres projets sont en cours qui utilisent les ontologies à des fins de représentation d'une connaissance, non plus sur l'image, mais sur le terrain. En collaboration avec les équipes de ces projets, il a été convenu d'envisager la perspective de tisser des liens entre la connaissance de terrain et la connaissance de l'image. Afin d'anticiper cette étape, OBOE a été choisie comme cadre de travail commun.

Avant d'être rattachées à OBOE, les visions de la structure et des traitements de l'image satellitaire étaient moins complexes. Ne reposant sur aucune ontologie de plus haut niveau d'abstraction, elles ne contenaient ni les observations, ni les mesures d'OBOE. Ce type d'ontologies était plus simple à utiliser et aurait probablement joué un rôle bénéfique sur les performances. Néanmoins, il aurait défavorisé le partage de l'information qui est un enjeu crucial de la formalisation des conceptualisations et des connaissances.

5.1.3 Objet de la conceptualisation de référence

Dans la proposition d'approche, la conceptualisation de référence est une base de départ utilisée pour la description de deux autres types de connaissance (FIGURE 5.5).

La première connaissance relève de l'expertise humaine. Elle dérive de la conceptualisation de référence par subsomption de concepts (FIGURE 5.7). Elle se rapporte à l'observateur, au sujet et c'est donc une connaissance plutôt subjective.

La seconde connaissance est liée à l'image. Elle prend également sa source dans la conceptualisation de référence, mais cette fois-ci par instanciation des concepts (FIGURE 5.8). Elle se rapporte à l'objet observé, ce qui lui confère une forte objectivité (dans le cadre de la conceptualisation de référence).

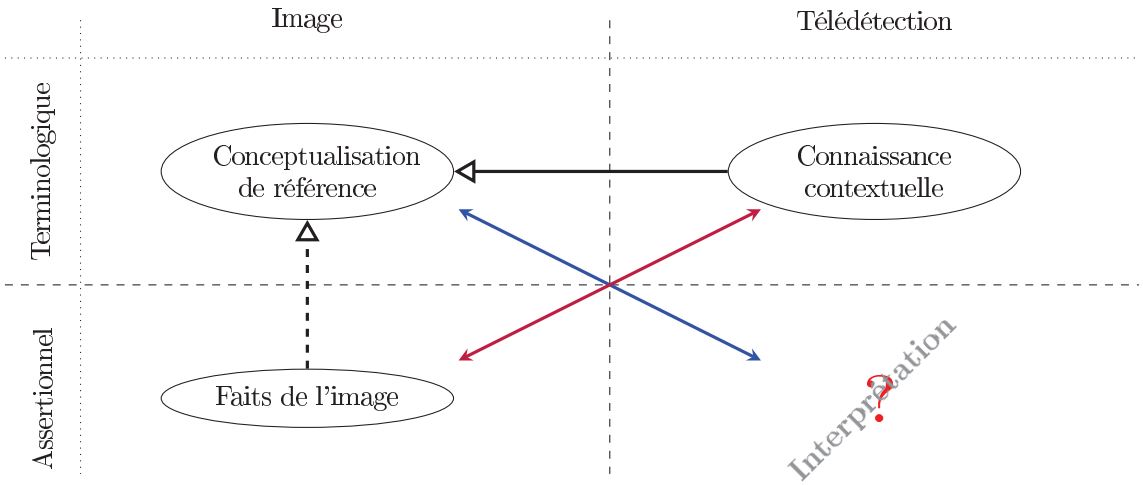


FIGURE 5.5 – La conceptualisation de référence est la base de la description du contenu de l'image par instanciation des concepts de référence et la base de la description de la connaissance humaine par spécialisation des concepts de référence.

Pour résumer (Conceptualisation de référence)**.** *La conceptualisation de référence est une ontologie (ou un ensemble d'ontologies) structurant et déterminant une certaine vision du monde par le biais de concepts et de rôles. Bien qu'elle puisse s'appliquer à un domaine bien particulier (par exemple, l'image satellitaire), elle est relativement indépendante de la diversité thématique opérationnelle (par exemple, les objectifs thématiques des experts).*

5.2 Connaissance contextuelle

Définition 68 (Connaissance contextuelle, adaptée d'après Falomir et al. (2011))**.** Une connaissance contextuelle représente des concepts d'un domaine dans le cadre spécifique d'une application. Leur définition peut s'appuyer sur une conceptualisation de référence, ce qui n'implique pas forcément qu'ils soient validés par l'ensemble de la communauté du domaine (cela est fonction des applications spécifiques de chacun).

Dans l'absolu, une connaissance contextuelle pourrait être valide dans une large part du domaine dans la mesure où elle est partagée par de nombreuses applications, de nombreux utilisateurs. Il faudrait pour cela qu'elle soit par exemple valide pour toutes les images satellitaires, quel que soit le capteur, la zone représentée, les objectifs et l'expérience de l'expert à l'origine de la connaissance.

Mais en pratique, tous ces facteurs varient d'une application à l'autre, réduisant d'autant le contexte de validité de la connaissance contextuelle. Telle connaissance contextuelle sera valide uniquement pour les images SPOT, telle autre uniquement pour des images de forêt tropicale, telle autre encore sera considérée valide uniquement par une équipe réduite de chercheurs du simple fait de l'originalité des objectifs qui guident leurs recherches. Rien n'empêche même que le contexte de validité se réduise à une seule personne paramétrant un unique traitement avec une connaissance spécifique à un objectif particulier.

Représentation d'un consensus, ou d'un avis personnel... il est probable que toute connaissance est un intermédiaire entre ces deux extrêmes.

La restriction du contexte d'utilisation d'une connaissance remet-elle en question la vocation des ontologies à jouer le rôle de vecteurs favorisant le partage de la connaissance ?

En ce qui concerne le partage entre acteurs humains, sans doute. Mais le destinataire de la connaissance peut également être un programme informatique sachant interpréter des règles logiques... c'est-à-dire un raisonneur !

De manière à mettre l'approche en pratique, une ontologie de connaissance contextuelle a été produite. Il s'agit de définitions subsumées par des concepts de la conceptualisation de référence (Figure 5.6). Elles sont construites à l'aide de règles basées sur les logiques de description et traduisant les caractères spécifiques permettant de déterminer la sémantique de télédétection à laquelle se réfère une entité de l'image : par exemple, un segment est-il représentatif d'une couverture végétale ?

$$VegetatedSegment \doteq Segment \sqcap \exists hasFeature.(FeatureNDVI \sqcap \exists Mean_{\{\geq 3,5\}})$$

Figure 5.6 – Construction type simplifiée d'un concept de la connaissance contextuelle. À droite de l'axiome d'équivalence, on trouve des concepts atomiques (en rouge), des rôles (en orangé) et des propriétés sur des domaines concrets (en jaune). Tous ces éléments sont issus de la conceptualisation de référence et combinés à l'aide de constructeurs logiques (en vert). À gauche de l'axiome d'équivalence, le concept atomique défini dans la connaissance contextuelle.

La connaissance contenue dans cette ontologie fait le lien entre le monde de l'image et la sémantique de télédétection. Néanmoins, les concepts qui y sont définis appartiennent toujours au monde de l'image. Il s'agit bien de caractériser un *segment image de végétation* et non pas de la végétation au sens propre... même si le vocabulaire couramment utilisé (mais trompeur) fait

souvent référence à l'entité de terrain elle-même. Ceci peut paraître du détail, mais revêt en fait une grande importance.

Nous avons parlé en première partie de cette distinction entre les entités physiques et leurs représentations numériques. Elle apparaît dans la littérature concernant non seulement l'information géographique (Castilla and Hay, 2008), mais également d'autres domaines comme les images médicales (Iakovidis et al., 2009).

Nous y voyons une triple justification.

1. Un argument de nature, avant toute chose. Une entité physique est composée de molécules alors que sa représentation numérique est une entité composée de nombres.

2. Un argument de caractéristiques intrinsèques, ensuite. Une entité physique et une entité numérique ne sont pas caractérisées de la même manière. On peut parler du *feuillage* d'une forêt... mais de l'*indice de végétation* de sa représentation dans une image.

3. Un argument de caractéristiques extrinsèques, enfin. Les entités physiques ne sont pas forcément caractérisées par les mêmes relations spatiales que les entités qui les représentent dans l'image. D'abord, parce que l'image est plane, et ensuite parce que toutes les entités physiques ne sont pas représentées dans une image. Prenons l'exemple d'une *forêt* séparée d'un *lac* par une *route*. La route peut très bien ne pas être représentée dans une image, si la résolution n'est pas suffisante. Une telle image montrera que les représentations de la forêt et du lac sont adjacentes, alors que les entités physiques ne le sont pas.

Hypothèse 1 (Hypothèse distinction des entités de l'image des entités de terrain)**.** On distingue les entités de l'image des entités physiques qu'elles représentent. Dans le cadre de la réduction du fossé sémantique, la connaissance contextuelle décrit des entités de l'image et non pas des entités géographiques de terrain.

Concrètement, la connaissance contextuelle contient des règles exprimant une *expérience humaine* à partir de concepts et de propriétés empruntés à la conceptualisation de référence qui représente l'image. Il s'agit en l'occurrence d'une ontologie de la télédétection.

Dans cette ontologie, chaque règle crée un nouveau concept porteur d'une sémantique de télédétection (par exemple, le concept de segment d'image représentant une zone végétalisée). Ces concepts sont des spécialisations d'une entité de l'image (FIGURE 5.7) mais dont le sens dépasse le strict monde de l'image pour intégrer une interprétation de télédétection.

Par conséquent, ce qui n'était qu'une instance de *Segment* dans le cadre de la conceptualisation de référence peut très bien être classé comme *segment de végétation* à la lumière de la connaissance contextuelle ainsi définie. Le segment n'a pas changé de nature et reste un segment d'image... Les règles exprimées dans la connaissance contextuelle appliquées aux caractéristiques du segment dans l'image permettent simplement de préciser qu'il s'agit là d'une instance de segments particuliers qui correspondent en fait à une surface végétalisée sur le terrain.

Dans un autre objectif, les raisonneurs peuvent aussi être appliqués à la conceptualisation de référence et à la connaissance contextuelle de manière à valider la consistance des terminologies.

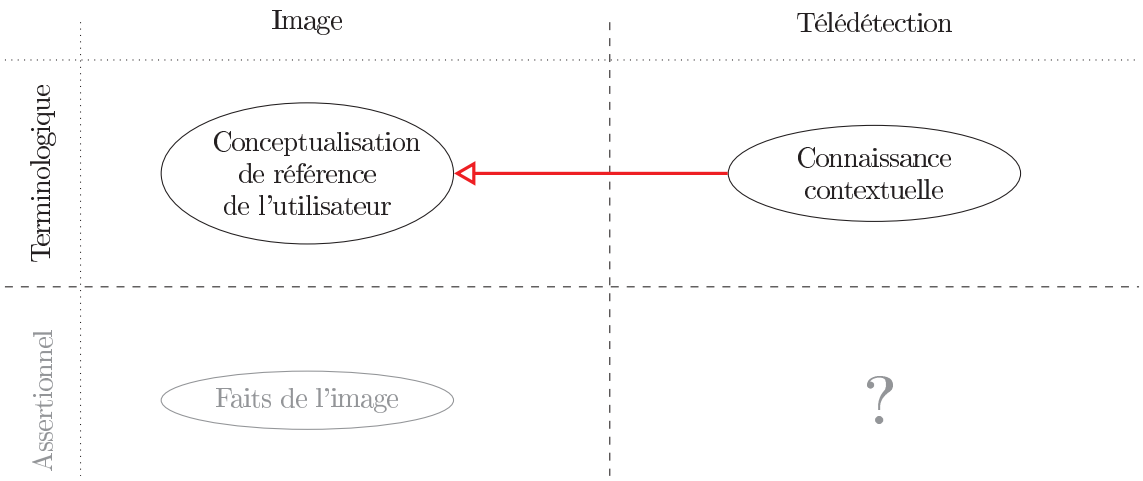


FIGURE 5.7 – Description de la connaissance contextuelle à partir de la conceptualisation de référence. Il s'agit principalement de spécialiser les concepts de la conceptualisation de référence pour chaque concept de télédétection que l'on souhaite représenter.

Pour résumer (Connaissance contextuelle)**.** *La description de la connaissance contextuelle se base sur une conceptualisation de référence pour exprimer le contenu d'une expérience acquise par un expert ou une communauté d'experts en télédétection, de manière subjective (en fonction de leurs objectifs thématiques).*

5.3 Peuplement par la description de l'image

Conceptualisation de référence d'une part et connaissance contextuelle d'autre part, interviennent principalement au niveau terminologique. Il s'agit d'axiomes c'est-à-dire d'énoncés supposés vrais dans les limites de leur contexte de validité. Les faits du monde sont eux représentés par des individus instances des concepts. Mis à part les traitements évoqués plus haut, les individus sont peu utilisés au niveau terminologique, sauf via le constructeur d'énumérations $\mathcal{O}$[1].

Le sujet de thèse concerne pourtant bien des entités concrètes du monde : des images, des segments d'image, des pixels... C'est bien de l'interprétation sémantique de ces entités concrètes qu'il s'agit. Là encore, les raisonneurs ont un rôle à jouer.

Mais avant de travailler sur une représentation du monde, il faut la construire, c'est-à-dire peupler l'ontologie d'individus représentant chacun un élément d'un domaine d'interprétation.

Le domaine d'interprétation est constitué des entités identifiables dans une image satellitaire. Un individu est créé dans l'ontologie pour chaque entité de l'image que l'on souhaite représenter (FIGURE 5.8). Les individus sont typés par des concepts de la conceptualisation de référence, en fonction de l'entité qu'ils représentent. Les individus représentant des pixels sont instances du concept de *pixel*, ceux représentant des segments sont typés par le concept de *segment* etc. de manière à ce que l'interprétation constitue un modèle de la base de connaissances ainsi élaborée.

Après la description de la conceptualisation du domaine de départ (monde de l'image) et la description de l'expertise, la description de l'image ainsi obtenue est la troisième pièce de l'approche.

[1] Il s'agit du constructeur recensé dans la TABLE 3.5, page 74.

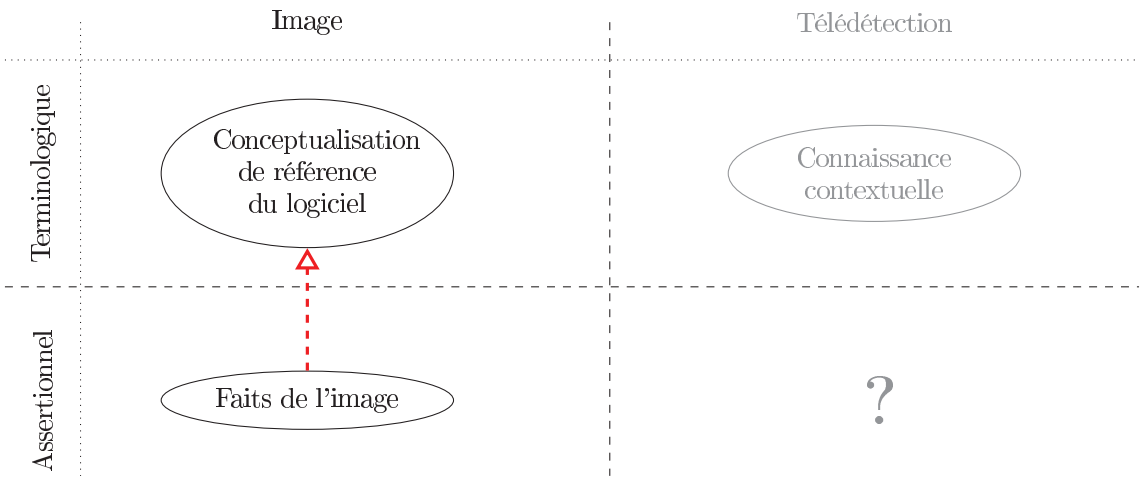


FIGURE 5.8 – Description des faits de l'image à partir de la conceptualisation de référence. Il s'agit principalement d'instancier les concepts de la conceptualisation de référence pour chaque entité de l'image que l'on souhaite représenter.

Pour résumer (Faits de l'image). *La description des faits de l'image se base sur une conceptualisation de référence pour exprimer le contenu d'une information produite par les traitements d'image en télédétection, de manière objective (indépendamment des objectifs thématiques des experts).*

5.4 Interprétation sémantique de l'image

Dans la mesure où la brique logicielle utilise une conceptualisation de référence (pour décrire l'image) et l'expert en utilise une autre (pour décrire la connaissance), alors la conceptualisation de référence désignée dans cette section désigne la réunion de :

- la conceptualisation de référence utilisée pour la description de l'image ;
- la conceptualisation de référence utilisée pour la description de la connaissance experte ;
- la correspondance entre les deux conceptualisations de référence.

5.4.1 Opération d'union sur les bases de connaissance

Définition 69 (Union sur les bases de connaissances). On remarque que l'union (au sens de l'union ensembliste) de deux bases de connaissances produit une base de connaissances et on note $\star$ cette loi de composition interne.

$\mathcal{C} = \mathcal{A} \star \mathcal{B} \iff \mathcal{C} = \mathcal{A} \cup \mathcal{B}$, avec $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ des bases de connaissances.

5.4.2 Analogie avec la règle de trois et le produit en croix

La simple description des faits, rattachée à une conceptualisation, confère déjà à l'image un certain degré de sémantique. En télédétection cela ne suffit pas toutefois, car l'interprétation attendue fait référence à une sémantique de terrain. En s'appuyant sur un formalisme logique, la connaissance contextuelle décrit précisément et explicitement, les règles utilisées intuitivement par l'expert dans cette opération sémantique.

Les illustrations proposées (FIGURE 5.7 et FIGURE 5.8) rappellent la structure d'un produit en croix. Selon cette vision des choses, l'inconnue se situerait au niveau assertionnel dans le domaine de la télédétection. Il s'agit donc bien de l'interprétation des faits de l'image dans une sémantique de télédétection (FIGURE 5.9).

Qu'obtient-on en procédant par analogie avec le produit en croix, avec la loi $\star$? Si l'on appelle $\mathcal{C}$ la conceptualisation de référence, $\mathcal{J}$ l'interprétation recherchée de l'image, $\mathcal{D}_\mathcal{K}$ la description de la connaissance experte et $\mathcal{D}_\mathcal{F}$ la description des faits de l'image, on pose l'égalité suivante :

$$\mathcal{C} \star \mathcal{J} = \mathcal{D}_\mathcal{K} \star \mathcal{D}_\mathcal{F} \tag{5.1}$$

Mais la description de la connaissance experte et la description de l'image ont toutes deux été élaborées à partir de la conceptualisation de référence et elles en contiennent les axiomes. On peut donc extraire $\mathcal{K}$ les axiomes de la connaissance experte et $\mathcal{F}$ les assertions sur les faits de l'image.

$$\begin{aligned}
\mathcal{C} \star \mathcal{J} &= \mathcal{D}_\mathcal{K} \star \mathcal{D}_\mathcal{F} && (5.2)\\
&= (\mathcal{C} \star \mathcal{K}) \star (\mathcal{C} \star \mathcal{F}) && (5.3)\\
&= (\mathcal{K} \star \mathcal{C}) \star (\mathcal{C} \star \mathcal{F}) && (5.4)\\
&= \mathcal{K} \star (\mathcal{C} \star \mathcal{C}) \star \mathcal{F} && (5.5)\\
&= \mathcal{K} \star \mathcal{C} \star \mathcal{F} && (5.6)
\end{aligned}$$

Mais $\star$ obéit aux mêmes règles que $\cup$. D'où, on a en particulier :

$$\mathcal{J} \subseteq \mathcal{C} \star \mathcal{J} \tag{5.7}$$

Et donc, en combinant (5.6) et (5.7) :

$$\mathcal{J} \subseteq \mathcal{K} \star \mathcal{C} \star \mathcal{F} \tag{5.8}$$

Ainsi, il ressortirait de ce schéma que la combinaison de la conceptualisation de référence, de la connaissance contextuelle et des faits de l'image contiendrait toute la connaissance nécessaire à l'interprétation des faits de l'image (FIGURE 5.9).

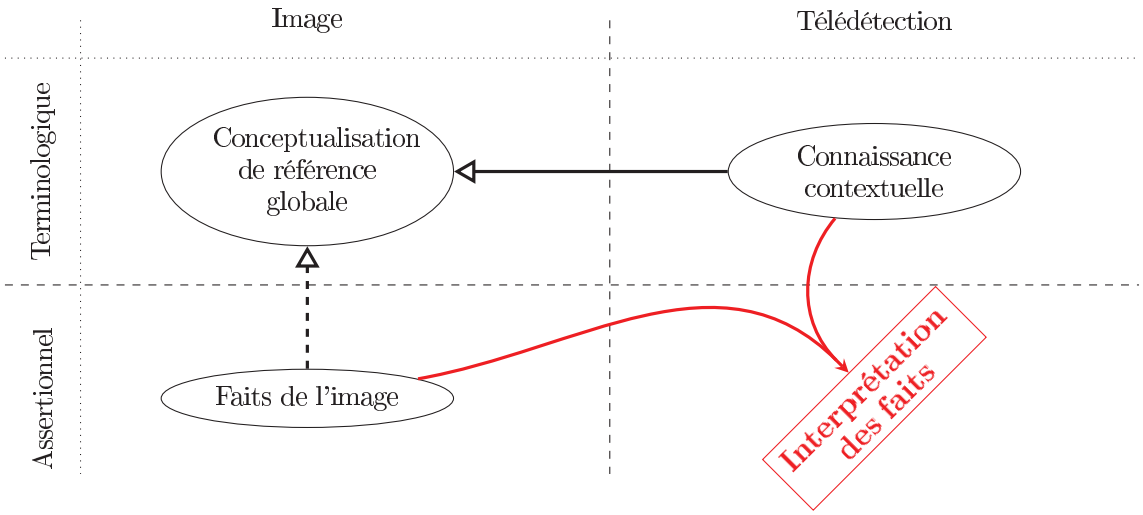


FIGURE 5.9 – Combinaison des différentes descriptions nécessaires à l'interprétation sémantique des faits de l'image.

Puisque le projet CARTAM-SAT se place dans la perspective de long terme d'une prise en charge automatique des images du satellite aux acteurs, il faut envisager d'appliquer ce principe dans un but d'interprétation automatique. On peut en conclure qu'un raisonneur doté des trois

connaissances qui ont été examinées successivement dans ce chapitre, devrait être en mesure d'inférer la classification des faits de l'image dans des classes correspondant aux concepts de télédétection explicités[1].

Quoiqu'intuitive, l'analogie avec le produit en croix illustre bien le principe d'interprétation sémantique proposé, étroitement liée à la conceptualisation de référence d'une part et aux deux descriptions qui en dérivent d'autre part.

[1] Sous la réserve technique de la compatibilité du raisonneur avec la logique de description utilisée.

6

Planification de la mise en œuvre

FORMULER LE PRINCIPE de l'approche mérite d'être complété par des considérations complémentaires liées aux questions de mise en œuvre. L'un des objectifs de ce chapitre consiste à se pencher sur les choix qui ont guidé la délimitation des segments dans l'image, en ce qui concerne l'application d'approches orientées objet. L'autre but est d'examiner, en fonction des choix techniques envisageables, les différents degrés de modularisation et d'intégration de la connaissance à l'algorithme.

Sommaire

6.1 Données en entrée **110**
6.1.1 Quelles données ? 110
6.1.2 Quelle bonne segmentation ? 110
6.2 Que faire de la connaissance en sortie de la segmentation ? **113**
6.2.1 Décrire la connaissance utilisée pour la segmentation 113
6.2.2 Ignorer la connaissance utilisée pour la segmentation 114
6.3 Intégration et modularité de la connaissance **116**
6.3.1 La source de la connaissance experte 116
6.3.2 La place de la connaissance par rapport aux traitements 117
6.4 Réflexions sur la modularisation de la connaissance **121**
6.4.1 Pourquoi modulariser la connaissance ? 121
6.4.2 Comment modulariser la connaissance ? 122
6.5 Réflexions sur la représentation des relations spatiales **122**
6.5.1 Représentation des relations topologiques 122
6.5.2 Interprétation de l'adjacence dans une image matricielle avec RCC8. . 123

6.1 Données en entrée

Nous allons explorer dans la mise en œuvre, plusieurs déclinaisons appliquées du principe formulé dans le chapitre précédent. Certaines suivent un paradigme d'analyse des pixels, d'autre un paradigme d'analyse d'objets. Toutes supposent que l'image utilisée en entrée est calibrée de manière à ce que les valeurs correspondent à des grandeurs physiques. Les approches par objets supposent une bonne segmentation telle que définie par Castilla and Hay (2008) (Définition 42 page 53). Se pose alors le problème du choix d'une méthode de segmentation.

6.1.1 Quelles données ?

6.1.1.1 Choix de la zone d'étude et du capteur

Bien que le projet CARTAM-SAT soit plus particulièrement dédié à l'environnement Guyanais et aux images SPOT, le travail de recherche sur une approche générique autorise de travailler sur des données différentes pour des raisons pratiques. Une fois formulé, le principe peut être mis en œuvre sur des données de sources variées, à condition que le contexte de validité de la connaissance corresponde à l'image utilisée.

Alors que des incertitudes ont longtemps plané sur la calibration des images SPOT[1], le travail collaboratif avec d'autres projets centrés sur les ontologies en télédétection a facilité l'accès à des images Landsat 5 dont la calibration radiométrique est fiable. La principale ressource utilisée ensuite est un extrait de prise de vue Landsat 5 du 29 juin 2010, des alentours de la ville de Santarém au Brésil (FIGURE 6.1).

Hypothèse 2 (Hypothèse de généricité de l'approche)**.** Le choix de la source de données n'affecte pas la pertinence de l'approche, pourvu qu'elle soit compatible avec le contexte de validité de connaissance formalisée.

6.1.1.2 Prétraitements requis

Nous avons vu plus haut que les pixels d'une image brute ne représentent directement aucune grandeur physique. Ce sont de simples comptes numériques qui ne peuvent être transformés en mesures physiques qu'à l'aide des métadonnées de l'image. C'est cette opération que l'on a appelée « calibration radiométrique » de l'image.

Or, nous avons vu par ailleurs que c'est principalement la signature spectrale qui permet de caractériser la nature des entités représentées sur une image satellitaire.

Afin d'utiliser la connaissance experte basée sur les signatures spectrales, il importe avant toute chose de travailler sur des images satellitaires calibrées. La FIGURE 6.1 représente la zone étudiée dans les longueurs d'onde visibles. Cette image a été calibrée en réflectance et température (selon le canal considéré) par un expert télédétecteur et géographe de l'unité.

Hypothèse 3 (Hypothèse de calibration de l'image en réflectance et température)**.** L'image en entrée est un extrait d'image multispectrale calibrée en réflectance et température (selon le canal considéré).

6.1.2 Quelle bonne segmentation ?

La sémantique physique élémentaire des pixels obtenue par calibration radiométrique est requise quel que soit le type d'analyse. Outre la diminution du nombre d'entités décrites (il y a davantage de pixels que de groupes disjoints de pixels) une approche objet permet de considérer de nouvelles

[1] L'un des deux paramètres essentiels à la calibration radiométrique (l'*offset*) est nul dans toutes les métadonnées examinées. L'interprétation de cette valeur comme fiable ou erronée a fait l'objet de débats à l'issue incertaine.

FIGURE 6.1 – L'extrait (825 × 666 pixels) de l'image Landsat 5 utilisée dans le développement d'une approche intégrée, dans une combinaison ici des longueurs d'onde visibles.

caractéristiques des entités, comme des indices relatifs à leur forme. Mais l'analyse orientée objet requiert un autre traitement préalable : la délimitation et l'identification de segments.

Dans le cadre d'une analyse orientée objet, on travaillera dans l'hypothèse suivante :

Hypothèse 4 (Hypothèse de bonne segmentation)**.** On se place en aval d'une approche orienté objet. Les segments existent comme point de départ de l'approche. Il n'est pas question de la méthode qui les a produit. On suppose simplement que les segments sont significatifs et qu'ils correspondent à des objets de l'image.

Il faut donc s'arrêter un instant sur la question de la segmentation. Il n'est pas question de rechercher ici l'algorithme miracle, dont nous avons vu qu'il est probablement inaccessible du simple fait que la segmentation est soumise aux variations d'une connaissance consubstantielle à son exécution. Il s'agit simplement de répondre à une question technique circonscrite à la mise en œuvre dans le cadre de la thèse : *Comment travailler sur des segments sémantiquement cohérents ?*

6.1.2.1 Classification « de BARALDI »

On l'a vu, l'ajustement et la compréhension des paramètres des algorithmes de segmentation populaires n'est pas chose aisée et ces programmes ne produisent pas forcément de résultats pleinement reproductibles (voir l'impact du contexte alentour d'une zone sur les résultats de l'algorithme *mean shift* en FIGURE 2.13, page 56).

Hypothèse 5 (Hypothèse de définition intrinsèque des segments)**.** On suppose que la définition d'un objet de l'image ne dépend pas du contexte autour de lui sur l'image. Deux algorithmes différents peuvent identifier des objets différents sur la même image. Par contre, un même algorithme de segmentation à paramètres constants doit produire des objets identiques sur une zone identique (voir le contre-exemple en FIGURE 2.13, page 56).

Le moyen le plus simple pour s'affranchir de telles variabilités, est d'opérer une préclassification de chaque pixel dans une classe d'occupation du sol, puis de réaliser une segmentation par seuillage de pixels sur la classification.

Les travaux de Baraldi et al. (2006) proposent justement une famille d'algorithmes de classification de cette sorte. Il s'agit d'arbres de décision appuyés sur des caractéristiques spectrales

complexes. *Orfeo Toolbox* inclut une implémentation de l'un d'entre eux, applicable aux images Landsat 5 ou Landsat 7 calibrées. Elle produit une image sur laquelle la valeur de chaque pixel correspond à la classe qui lui est attribuée (FIGURE 6.2).

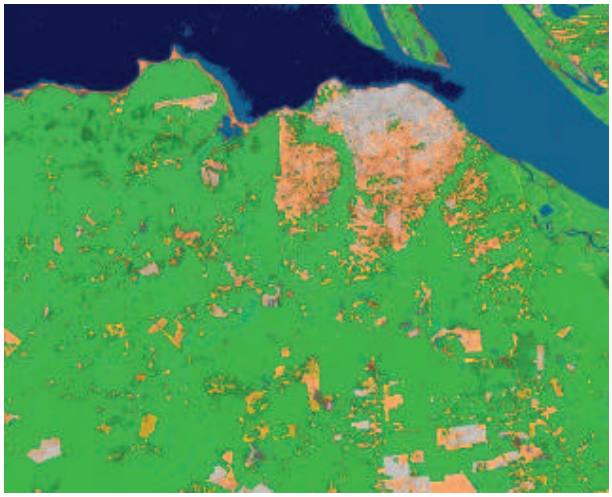


FIGURE 6.2 – Le résultat coloré de la préclassification de BARALDI appliquée à l'image présentée en FIGURE 6.1. Ici, chaque couleur représente une préclasse particulière.

Il reste alors à effectuer la segmentation par seuillage de manière à ce que chaque groupe de pixels connexes de même valeur dans l'image de préclassification de BARALDI constitue un segment (FIGURE 6.3). Par construction, cette méthode de segmentation produit des segments dont la cohérence sémantique découle de la préclassification de BARALDI.

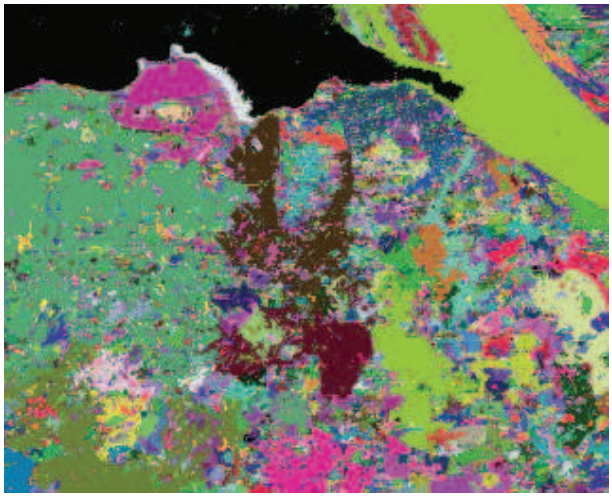


FIGURE 6.3 – Le résultat coloré de la segmentation basée sur la préclassification de BARALDI présentée en FIGURE 6.2. Ici, chaque région colorée représente un segment d'image.

6.1.2.2 Discussion

Pour décrire l'image et séparer la connaissance de l'algorithme, on commence par utiliser un algorithme incluant lui-même clairement une connaissance en télédétection. Ceci est paradoxal, mais est-ce rédhibitoire ?

Ce paradoxe est lié à un autre. Nous avons déjà cité Casati et al. (1998) pour lesquels la frontière d'une entité tient un rôle important dans la construction de son identité, par la distinction de ce qui

n'est pas d'elle. Les rapports de l'approche que nous avons développée vis-à-vis de la segmentation sont complexes et il est sans doute préférable de ne pas s'y aventurer au-delà du nécessaire, au risque de s'y perdre. Remarquons simplement qu'elle en est *dépendante* dans le sens où la sémantique d'un objet dépend de la définition de ses contours. Mais elle en est aussi *indépendante* dans le sens où elle peut être appliquée à la suite d'une segmentation quelconque, pourvu qu'elle produise des objets sémantiquement cohérents.

En fait, ce paradoxe n'est qu'apparent car il compare deux critères différents de l'approche : d'une part sa structure et d'autre part ses résultats en sortie en fonction des données d'entrée. L'approche est structurellement indépendante de la segmentation : elle ne requiert qu'une bonne segmentation, quel que soit l'algorithme utilisé pour la produire. Mais en ce qui concerne les résultats, comme il y aurait sans doute autant de bonnes segmentations idéales que d'experts en télédétection, le fruit de l'interprétation d'une image reste lui dépendant de la segmentation : si les contours des objets n'ont pas de sens, alors il est vain de rechercher une sémantique de télédétection qu'ils n'ont pas.

La segmentation basée sur une préclassification de BARALDI est simplement un moyen de valider l'hypothèse d'une bonne segmentation. On peut la remplacer par une autre bonne segmentation sans que le principe de l'approche soit altéré. Pour des raisons pratiques, c'est cette segmentation précise qui a été utilisée et il se trouve qu'elle se base sur une connaissance experte directement liée à la télédétection (ce sont des classes d'occupation du sol et non des variables mathématiques) et incluse dans l'algorithme (ce ne sont pas des paramètres de l'algorithme). Mais y a-t-il un moyen d'obtenir une segmentation parfaite sans connaissance experte ? En pratique, rappelons que l'ajustement des paramètres des autres algorithmes par le télédétecteur procédant par essai/erreur en fonction de son expérience incite à répondre par la négative. De même, la proposition de segmentation génétique et supervisée de Derivaux et al. (2010) intègre une connaissance experte, quoique de façon différente.

L'utilisation d'une préclassification experte par hypothèse en amont de l'approche n'est donc pas fondamentalement dérangeante. Il reste que la connaissance experte utilisée en entrée de la segmentation est toujours présente à la sortie. Comment faut-il la considérer ? Plusieurs solutions sont envisageables.

6.2 Que faire de la connaissance en sortie de la segmentation ?

Du fait du choix particulier concernant la segmentation dans le cadre de la préparation de la mise en œuvre, une connaissance experte en télédétection est disponible dès avant la description ontologique de l'image. Deux perspectives s'ouvrent dès lors : utiliser cette connaissance, ou bien l'ignorer.

6.2.1 Décrire la connaissance utilisée pour la segmentation

De la même manière que l'on a proposé de décrire les objets de l'image par leurs caractéristiques de bas niveau (radiométrie, texture...), il serait également possible d'en donner une description intégrant la connaissance experte en sortie de la segmentation.

À la sortie de la segmentation suivant la préclassification de BARALDI, chaque objet est constitué de pixels connexes dont la préclasse de BARALDI est identique. Rien n'empêche de décrire cette préclasse pour chaque objet, d'une manière similaire aux autres caractéristiques. Cette description de plus haut niveau[1] résumerait en quelque sorte la description radiométrique de bas niveau[2]. D'une certaine manière, cette dernière, plus volumineuse pourrait même devenir superflue et encombrante.

La description de la préclasse des objets dans les faits de l'image a été implémentée, mais n'a pas été poussée plus avant. Que faire de cette description ? Qu'elle soit ou non accompagnée d'une connaissance de bas niveau sur l'image, deux voies se présentent :

[1] Dans le sens où elle s'éloigne de caractéristiques de l'image pour se rapprocher de la télédétection.

[2] Puisque ce sont ces caractéristiques qui ont été utilisées dans l'algorithme de préclassification pour déterminer les préclasses de chaque pixel, les préclasses peuvent être considérées comme une synthèse de plusieurs caractéristiques radiométriques.

La première est simple. La description des préclasses des objets constitue une annotation sémantique achevée de l'image satellitaire. « Achevée », puisque du fait de cette annotation, les entités de l'image sont porteuse d'une sémantique liée à la télédétection (une occupation du sol). Cette voie peut donc être considérée comme la possibilité d'une annotation sémantique automatique des images satellitaires et pourrait être exploitée soit par des algorithmes spécifiques à la manière de Srikanth et al. (2005), soit par des raisonneurs. Voilà une piste intéressante, mais qui nous éloigne de la réduction du fossé sémantique et de l'interprétation sémantique pour nous rapprocher de questions qui abordent la fouille de données...

La seconde voie est plus exigeante. La description issue d'une connaissance sur les pixels, n'est pas considérée comme définitive car elle peut être enrichie par une connaissance sur les objets : des textures, la forme, les relations spatiales. Les faits de l'image relatifs à la préclasse des objets peuvent alors être traités selon le principe développé autour du raisonnement alimenté par la conceptualisation de référence et une connaissance contextuelle. La seule différence entre les descriptions des préclasses et les autres (radiométrie, forme, texture...) réside en l'aspect *qualitatif* des premières et l'aspect *quantitatif* des secondes. Les faits relatifs aux préclasses ne sont plus aussi bruts, bas niveau et objectifs qu'une description *quantitative* utilisant directement la radiométrie des objets. L'algorithme de préclassification joue alors le rôle du modèle qualitatif utilisé par Falomir et al. (2011) : c'est déjà une interprétation. Mais cette partie de l'interprétation n'utilise pas une connaissance formalisée en ontologie et limite d'autant le rôle auquel elles peuvent prétendre dans la réduction du fossé sémantique. Or, nous souhaitons, nous, capitaliser au mieux la connaissance au moyen des ontologies.

L'organisation des travaux de thèse nécessite de faire des choix. Nous avons choisi les voies à explorer en fonction des objectifs du projet CARTAM-SAT, des précédents travaux étudiés dans l'état de l'art bibliographique et des directions de recherche qui semblaient offrir le plus de perspectives de développement au sens du nombre de possibilités. La recherche de la performance, qui aurait dû peser en faveur de la description des préclasses, n'a pas été l'argument décisif. Avec les ontologies, la bataille est déjà perdue sur ce plan. Il faut donc pousser d'autres arguments, peut-être moins séduisants mais plus fondamentaux.

6.2.2 Ignorer la connaissance utilisée pour la segmentation

Il est également possible d'ignorer la connaissance experte utilisée dans la préclassification. Cette hypothèse permet de s'abstraire du choix circonstantiel de segmentation. Notre démarche souhaite en effet privilégier une certaine généricité : elle doit être reproductible, quelle que soit la bonne segmentation, même si elle n'est pas accompagnée en sortie d'une première connaissance experte de télédétection sur les objets.

Dans ce cas, on considère le résultat de la segmentation en « oubliant » la préclasse des pixels regroupés en objets : on n'en donne pas description dans les faits de l'image. Les objets ont une cohérence, mais sont complètement dépourvus de la moindre sémantique de télédétection. Reste à déterminer cette dernière uniquement à l'aide d'une description de bas niveau[1] des faits de l'image. C'est la voie qui a été privilégiée dans l'approche proposée, car c'est celle qui rapproche la connaissance au plus près de l'image.

L'abandon de la connaissance utilisée pour la préclassification pourrait être quelque peu frustrant, mais l'ambition globale de notre approche est à la fois plus modeste et plus profonde. Plus modeste en télédétection, car elle ne prétend pas décrire l'image avec autant de précision qu'une préclassification en 46 classes. Plus profonde sur la manière d'utiliser l'information, car elle cherche à exploiter le plus largement possible les ontologies et le raisonnement logique pour trouver ou retrouver une sémantique de télédétection des entités de l'image. C'est aussi en poussant les principes de la proposition que son potentiel et ses limites devraient apparaître manifestement.

On le voit, le potentiel des ontologies pour l'exploitation des images satellitaires est varié (FIGURE 6.4). Encore sommes-nous resté à proximité immédiate de l'image, sans déborder vers les questions de terrain.

[1]La description se limite uniquement à des concepts de l'image (indices, textures, formes...) sans faire appel à des notions plus complexes de télédétection.

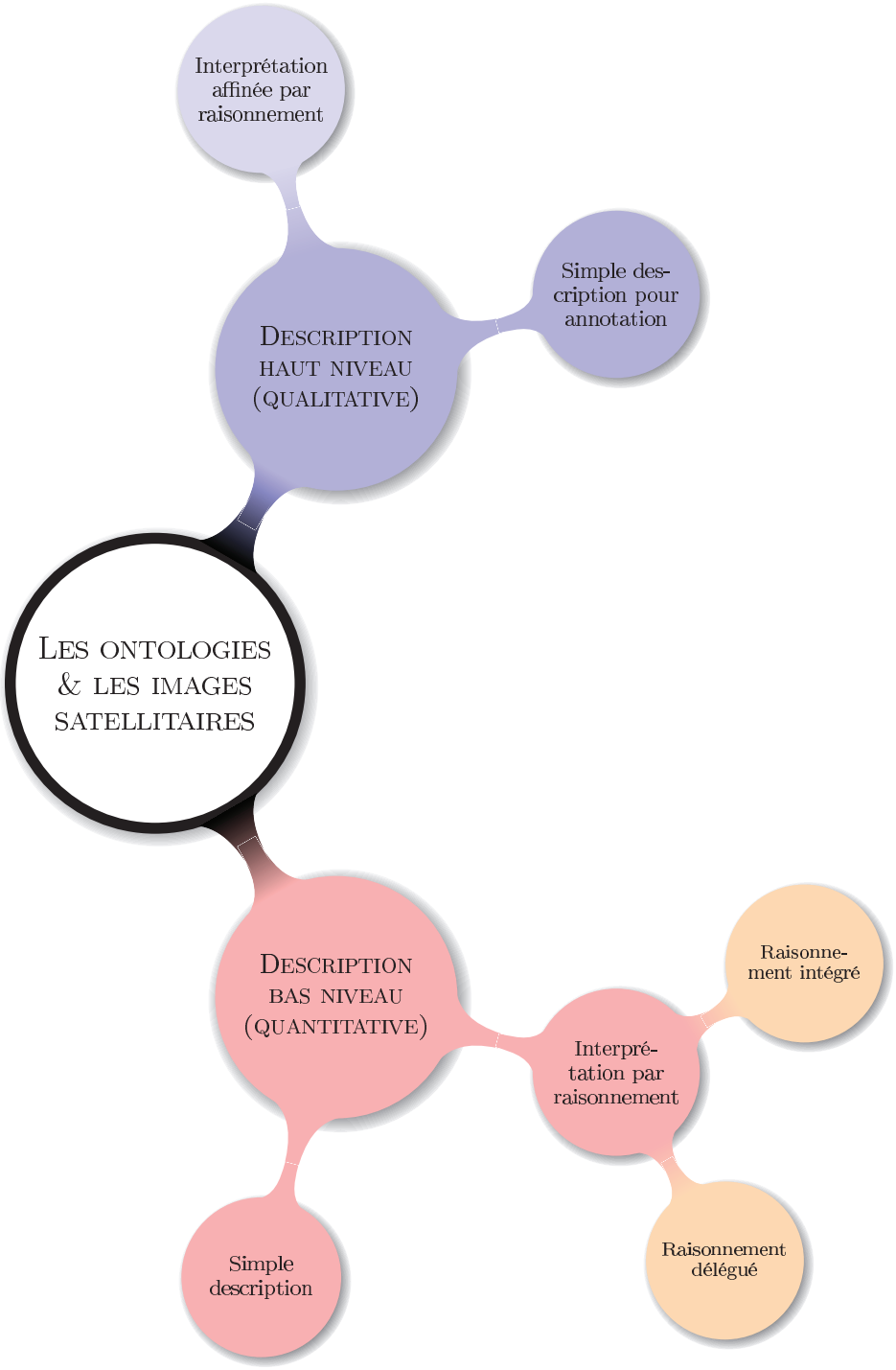


Figure 6.4 – Différentes pistes d'utilisation des ontologies dans les images satellitaires. En privilégiant une description ontologique quantitative (en rouge) par rapport à une description qualitative (en bleu) on rapproche d'autant la connaissance du monde de l'image. La piste qualitative offre deux possibilités : décrire simplement une interprétation classique, ou bien l'affiner par raisonnement automatique. De la même manière, une approche quantitative permet soit de décrire simplement le contenu de l'image, soit de raisonner sur cette description. C'est cette dernière piste que nous avons privilégiée. Des considérations plus techniques permettent ensuite d'envisager, soit un raisonnement assez fortement intégré au traitement d'image, soit un raisonnement séparé.

6.3 Intégration et modularité de la connaissance

6.3.1 La source de la connaissance experte

La question de la source de connaissance experte est à elle seule, comme la segmentation, un vaste problème. Nous l'avons vu, la connaissance contextuelle peut tout aussi bien représenter le consensus stable d'une communauté d'experts que la connaissance spécifique d'un spécialiste quant à un problème très particulier : cela fait partie du contexte de validité de la connaissance. L'origine de la connaissance peut donc varier : là ne réside pas vraiment la difficulté. En revanche, quelle forme la connaissance formalisée doit-elle prendre ? Car après tout, la connaissance incluse dans l'algorithme de préclassification de BARALDI que nous avons évoquée, est formalisée[1] !

Comme ce fut le cas vis-à-vis de la segmentation, il s'est agit d'échapper à un double piège concernant la formalisation de la connaissance experte. Premièrement, le sujet de thèse n'est pas de formaliser une connaissance en télédétection, qu'elle soit d'un expert ou d'une communauté d'experts. Deuxièmement, le développement du sujet nécessite la formalisation de quelques concepts de base afin de mettre le principe à l'épreuve de la pratique.

Trois possibilités se sont présentées au cours de la thèse, comme sources potentielles de connaissance experte.

6.3.1.1 Utilisation d'une connaissance formelle

La première réside en l'exploitation de l'algorithme de BARALDI implémenté dans la bibliothèque *Orfeo Toolbox*. Comme nous l'avons remarqué, il s'agit par définition d'une connaissance experte formalisée. Cela n'est pas son unique avantage, puisque son contexte de validité est parfaitement déterminé : il s'agit de l'interprétation spectrale d'A. BARALDI[2] sur les capteurs multispectraux des satellites Landsat 5 et Landsat 7. Cette connaissance se révèle toutefois complexe à représenter sous forme ontologique, en particulier parce qu'elle utilise des fonctions agrégatives sur les valeurs spectrales de chaque pixel à classifier. Or, l'intégration des fonctions agrégatives dans les ontologies n'est pas triviale et pose des problèmes de décidabilité.

Cela pose au passage le problème de la limite pragmatique entre la part d'information de l'image à représenter sous forme ontologique et la part à préserver du côté du traitement d'image. En fait, cette question s'est présentée bien plus tôt dans le développement de l'approche, mais elle est restée très discrète.

En effet, la donnée de l'image calibrée avec ses métadonnées est strictement radiométrique. Pourtant, dans la représentation que nous en proposons figurent à côté des valeurs de réflectance et de température, des valeurs d'indices et de textures qui ont été calculées par traitement d'image. Il y a donc eu un enrichissement sémantique de l'image par des grandeurs intermédiaires entre la radiométrie d'un côté et la sémantique de télédétection de l'autre. Cet enrichissement sémantique relève des opérations arithmétiques définies par les formules d'indices et de textures. Il semblait naturel alors de laisser cette tâche au traitement d'image et de ne représenter dans les ontologies que les opérations logiques. Il n'empêche que fondamentalement la question aurait pu se poser.

Si l'on souhaite utiliser l'algorithme de préclassification de BARALDI comme source de connaissance experte, il s'agit de suivre concrètement le chemin depuis une classe de la préclassification jusqu'à l'information de l'image (les valeurs spectrales d'un pixel). La classe elle-même est définie par la combinaison logique de plusieurs filtres, mais chaque filtre combine les valeurs des pixels avec des opérateurs arithmétiques, logiques et éventuellement des fonctions agrégatives. La solution pourrait donc être de recenser ces filtres comme autant d'indices élémentaires, exactement comme le sont les indices spectraux classiques. Mais il s'agit d'une tâche technique volumineuse dans le cadre d'une thèse. Il est souhaitable de s'orienter préférentiellement vers des activités de recherche plutôt que d'ingénierie.

[1] Bien que formalisée dans l'algorithme, cette connaissance n'est pas du tout une ontologie car elle ne fait appel à aucun concept. On peut également considérer que « prisonnière » du code source, sa spécification n'est pas particulièrement explicite.

[2] Cette interprétation, publiée (Baraldi et al., 2006), est donc relativement admise dans la communauté.

6.3.1.2 Utilisation de techniques d'apprentissage

Une deuxième possibilité serait de procéder par apprentissage. Différentes approches développées dans l'unité de recherche proposent d'expliciter la connaissance experte par apprentissage. Arvor et al. (2013b) extraient ainsi les caractéristiques permettant de discriminer au mieux des classes d'occupation du sol les unes par rapport aux autres. Bayoudh et al. (2012) proposent une approche inductive d'apprentissage de règles structurelles.

6.3.1.3 Élicitation de la connaissance experte

Reste une troisième possibilité : l'approche directe d'élicitation de la connaissance experte. Bien que de manière générale, l'élaboration d'une ontologie suive des étapes bien déterminées (Laporte, 2011), de nombreuses méthodes existent pour appréhender une expertise. Shadbolt and Burton (1995) en proposent une recension.

La discussion directe, qui semblait l'approche la plus pertinente à l'intérieur de l'unité de recherche présente néanmoins l'inconvénient d'un certain *quiproquo* entre l'informaticien et le télédétecteur concernant la formalisation et la désambiguïsation de la connaissance. Un moyen terme semblait donc bienvenu pour éviter la prolongation des discussions et atténuer les ambiguïtés. De manière à engager le contact avec les télédétecteurs, nous leur avons proposé de commencer par désigner individuellement quelques unes de leurs classes d'intérêts. Chaque classe devait ensuite faire l'objet d'une description dans un tableau de caractéristiques. Certaines colonnes spécifiaient des caractéristiques de l'image précises (indices spectraux...), d'autres laissaient davantage de liberté (descriptions qualitatives...) et enfin une colonne permettait de donner une description totalement libre de la classe d'intérêt avec la possibilité de faire appel à une description de terrain plutôt qu'à des caractéristiques de l'image. Il s'agissait principalement par ce biais de cerner les concepts manipulés par les télédétecteurs tout en les amenant à se poser la question de leur description plus ou moins formelle. Ces premières descriptions ayant en quelque sorte défriché la définition des concepts, elles auraient pu être approfondies ensuite par discussion directe. Malheureusement, cette initiative, peut-être trop austère et rébarbative, n'a pas obtenu l'écho espéré. Ce retour n'est cependant guère surprenant, puisque Shadbolt and Burton (1995) soulignaient déjà que l'élicitation d'une connaissance experte était une tâche extrêmement difficile.

Fort heureusement, cela n'a pas d'incidence sur le principe de l'approche. En attendant de les compléter par les résultats des approches par apprentissage ou par une étude spécialement dédiée à l'élicitation des concepts de télédétection, quelques concepts d'étude ont été construits à partir de références bibliographiques et techniques et de l'expérience acquise en cours de thèse.

6.3.2 La place de la connaissance par rapport aux traitements

Autant le choix du logiciel de traitement d'image a été arrêté très tôt dans le déroulement de la thèse, autant les outils de manipulation de la connaissance formalisée ont varié.

Nous présentons ici successivement la progression de l'approche sous l'aspect de l'intégration et de la modularité de la connaissance dans la proposition d'approche. La première proposition n'était pas du tout intégrée. La deuxième et la troisième ont permis, en se familiarisant à des outils spécifiques de mettre en œuvre deux types d'intégration modulaire, en séparant d'abord la connaissance de l'algorithme, puis du code source.

6.3.2.1 Traitement d'image et connaissance : approche non intégrée

CARTAM-SAT ambitionne la réalisation d'une chaîne de traitements intégrée, depuis la réception des données jusqu'à la diffusion de la connaissance. L'intégration est donc un aspect à prendre en compte. C'est en partie pour cette raison que la simple description de l'image dans un fichier OWL n'a pas été maintenue.

D'autre part, le logiciel *Soprano*[1] utilisé pour écrire la description en respectant les standards du web sémantique, ne supporte pas le passage à l'échelle exigé par des extraits d'images un tant soit

[1] http://soprano.sourceforge.net/.

peu volumineux. Il s'agit d'un programme libre[1] faisant partie intégrante du projet NEPOMUK[2] destiné à enrichir le bureau libre KDE de fonctionnalités sémantiques. Il n'est pas du tout adapté aux volumes de données courants en imagerie satellitaire.

Enfin, une fois la description de l'image réalisée, elle est exploitée par un raisonnement ontologique dont il s'agit de visualiser les résultats. L'analyse de l'image était réalisée dans un premier temps, avec un premier logiciel dédié : l'interface de programmation *Orfeo Toolbox*. La description ontologique relevait d'un deuxième temps s'achevant avec la sérialisation des faits dans des fichiers OWL à l'aide de *Soprano*. Toutefois, cette deuxième étape était intégrée à la première. Enfin, le raisonnement dépendait d'un troisième logiciel : l'éditeur d'ontologies *Protégé*. Ce dernier nécessitait l'importation manuelle des fichiers de description de l'image, étape qui n'était par conséquent pas du tout intégrée dans la chaîne de traitements (FIGURE 6.5). De plus, un éditeur d'ontologies ne permet pas de visualiser les résultats de l'interprétation sémantique des régions sous forme d'image.

Pour toutes ces raisons, la mise en œuvre a toujours tendu à favoriser une proposition intégrée de la connaissance au traitement d'image tout en s'appliquant à modulariser au mieux la connaissance experte.

6.3.2.2 La connaissance hors de l'algorithme

Nous avons cherché à intégrer le raisonnement directement à la suite du traitement d'image. Pour cela, c'est le raisonneur FaCT++ qui a été choisi (Tsarkov and Horrocks, 2006). Ce raisonneur, toujours activement développé, est une réécriture en C++ du raisonneur FaCT implémenté en LISP. Suivant les circonstances, FaCT++ se révèle plus ou moins performant que ses concurrents (cela dépend des ontologies), mais utiliser son implémentation native en C++ permet déjà d'éviter un goulet d'étranglement[3] (Bail et al., 2010). Comme les autres logiciels utilisés, il est publié sous licence libre[4].

Cette architecture intégrée qui prend en entrée une image satellitaire multispectrale calibrée, produit en sortie des masques correspondant aux segments classifiés dans les différents concepts de télédétection. Le résultat peut donc être considéré comme une collection d'images dotées d'une sémantique de télédétection.

Mais l'intégration forte de l'étape de raisonnement aux étapes de traitement d'image et de description a un prix. Le raisonneur doit avoir accès à la connaissance experte. Il faut pour cela décrire toutes les connaissances et les conceptualisations via l'interface de programmation du raisonneur.

La connaissance est décrite dans le programme sous une forme déclarative avant d'être utilisée par les algorithmes du raisonneur. L'intégration du raisonnement, tout en maintenant la séparation entre la connaissance et les structures algorithmiques, la confisque néanmoins à l'expert qui se voit obligé de se faire programmeur, s'il désire en garder la maîtrise.

6.3.2.3 La connaissance hors du code

Contrairement à une approche classique par arbre de décision, les ontologies ont contribué à séparer la connaissance de l'algorithme.

Que ce soit au moyen de la proposition élaborée avec *Soprano* ou avec FaCT++, l'expertise ne se trouve plus imbriquée dans une structure algorithmique, mais dans une base de connaissances séparée[5]. Inversement, on peut considérer que la composition d'un arbre de décision par l'expert (ainsi que le proposent des logiciels d'aide à la télédétection) maintient la connaissance dans l'algorithme (l'arbre de décision), bien qu'elle soit manipulée hors du code.

Il peut être intéressant d'isoler encore davantage la connaissance en la séparant non seulement de l'algorithme, mais également du code.

Il est tout-à-fait normal que l'approche intégrée alimente directement le raisonneur avec les faits de l'image, puisque c'était là l'objectif de cette option. De même, la conceptualisation de référence

[1] Sous Licence Publique Générale Limitée GNU 2.0 (http://sourceforge.net/projects/soprano/).

[2] http://userbase.kde.org/Nepomuk.

[3] Nous verrons en détaillant la mise en œuvre, que la bibliothèque *Orfeo Toolbox* utilisée pour les traitements d'image est implémentée en C++. Par conséquent, faire appel à des interfaces de programmation dans ce même langage permet de réaliser une forte intégration et d'éviter le passage d'un langage à l'autre qui peut être coûteux.

[4] Licence Publique Générale Limitée GNU (http://code.google.com/p/factplusplus/).

[5] Cette base de connaissances peut être explicitée dans un fichier (comme avec *Soprano*) ou dans le code source (comme avec FaCT++), elle n'en reste pas moins déclarative et distincte des structures algorithmiques.

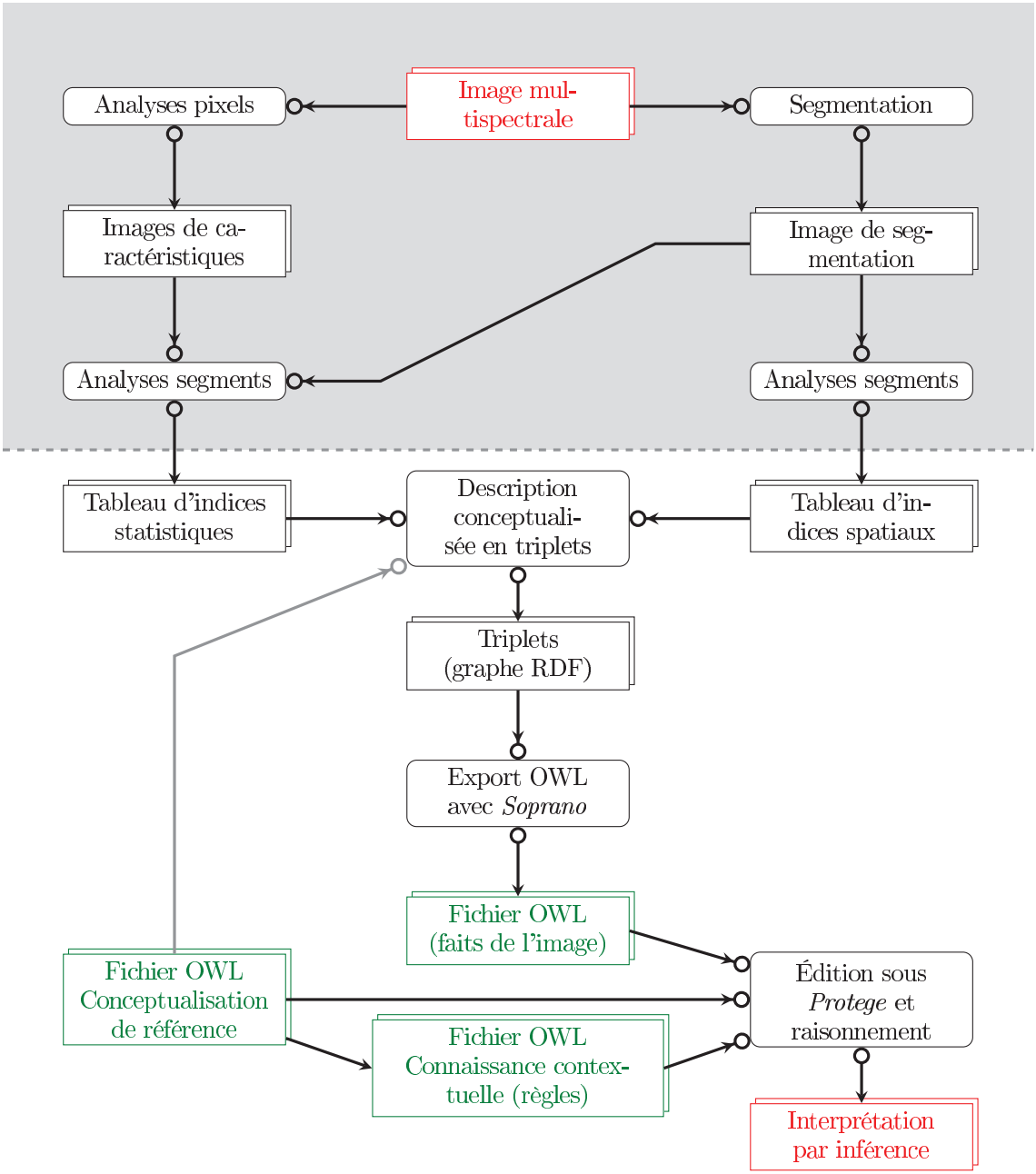


FIGURE 6.5 – Schéma de principe de la validation de l'approche théorique. Le processus part d'une image satellite et aboutit à sa description sémantique. La connaissance structurée (en vert) intervient en aval du traitement d'image. La description conceptualisée en triplets se réfère à la conceptualisation de référence, mais n'utilise pas le fichier OWL dans lequel elle est définie (d'où le lien grisé).

est définie dans le programme, puisque les faits de l'image s'appuient sur elle et qu'elle reflète en quelque sorte le point de vue de la brique logicielle logiciel[1] sur la structure et les traitements de l'image.

Il en va tout autrement de la connaissance contextuelle. Si l'on se contente simplement de la séparer de l'algorithme, elle reste définie dans le programme pour alimenter directement le raisonneur. Mais cette connaissance est censée avoir l'expert pour origine et non pas la brique

[1] En réalité, ce point de vue est celui du programmeur.

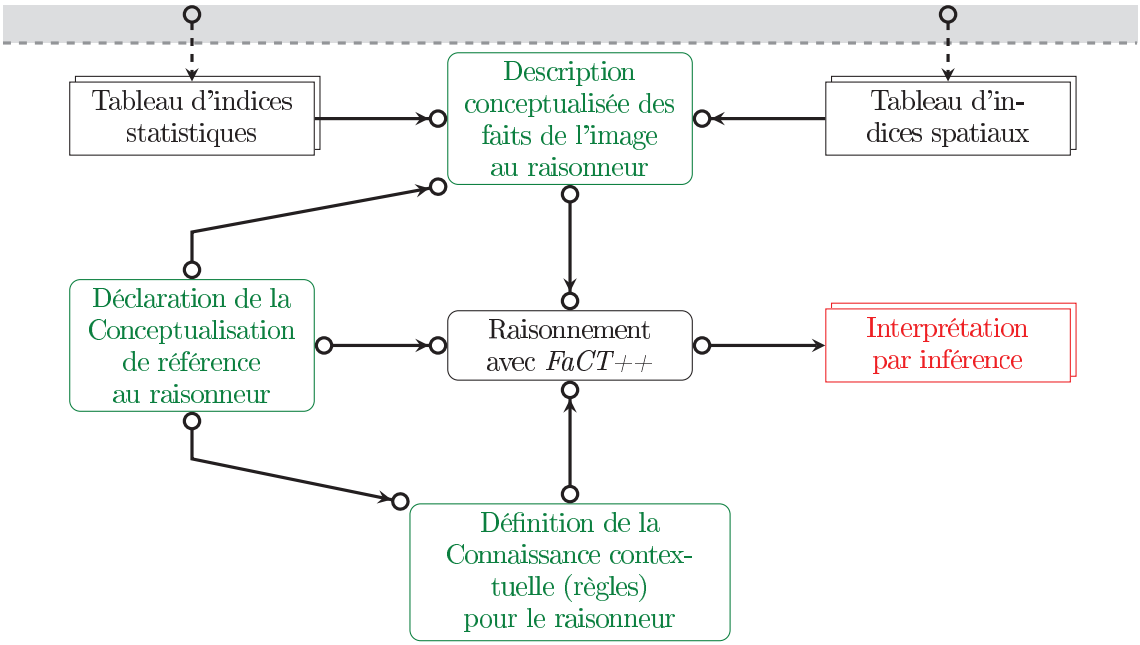


FIGURE 6.6 – Schéma de principe d'intégration du raisonnement (voir la FIGURE 6.5 pour la partie amont).

logicielle. Il serait donc logique qu'elle ne soit pas définie dans le code. Dans le cas contraire, le principe de l'approche n'en est pas affecté, mais la réalisation s'en trouve dissonante.

Mais FaCT++ n'est qu'un raisonneur. Il ne dispose pas d'une interface de programmation native lui permettant de lire des formats aux standards du web sémantique. Il est donc exclu de définir la connaissance experte à l'extérieur du programme, dans des fichiers dédiés et modifiables à volonté[1].

Pour utiliser ces fonctionnalités, il faut passer par l'interfaçage Java et OWL API... ce qui complique l'intégration du raisonnement[2]. De plus, l'intégration de modules logiciels écrits en langages différents est susceptible de poser des problèmes de performances.

Or, Levin et al. (2011) proposent de pallier le manque d'un parseur OWL pour FaCT++. Le logiciel proposé, owl_cpp, est libre[3]. En l'état de son développement, il ne permettait pas de lire un fichier contenant les règles définies par l'expertise, en particulier à cause de lacunes quant au support des intervalles de valeurs. Des échanges répétés avec l'auteur de cette bibliothèque à propos des différents problèmes rencontrés ont permis d'atteindre un niveau de développement suffisant à cet objectif.

Ce maillon supplémentaire complète l'approche (FIGURE 6.7). Il parachève et justifie la séparation de la connaissance et de l'algorithme en extrayant la connaissance experte du code même. Bien que n'apportant rien de fondamental, cette touche technique n'est pas négligeable car elle donne tout son sens au potentiel de partage des ontologies. En effet, pour être structurée et formalisée, la connaissance prisonnière du programme ne peut être facilement partagée, communiquée, échangée, modifiée. Raisonnement automatique ou arbre de décision, la différence apparente est maigre si l'utilisateur n'a pas la main sur les ontologies censées représenter sa connaissance.

owl_cpp permet donc de compléter le potentiel des ontologies dans la proposition d'interprétation des images satellitaires par raisonnement.

[1] À moins d'utiliser une syntaxe fonctionnelle, proche du LISP et héritée du raisonneur FaCT.

[2] Signalons en outre que la thèse a nécessité l'adaptation à des outils de domaines très différents du traitement d'images satellites au raisonnement sur les logiques de description. L'adaptation technique est un problème de base, mais il prend beaucoup de temps. S'attaquer à la combinaison de Java et C++ semblait dépasser le cadre de la thèse.

[3] Licence Logicielle Boost 1.0 (http://sourceforge.net/projects/owl-cpp/).

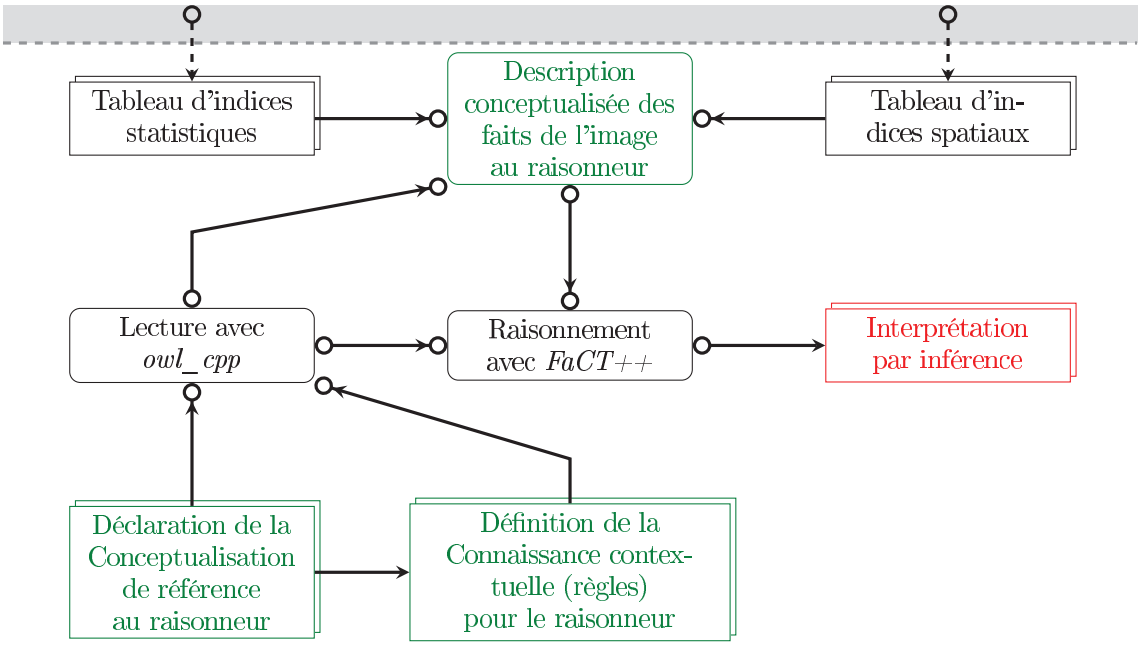


FIGURE 6.7 – Nouveau schéma de principe d'intégration du raisonnement (voir la FIGURE 6.5 pour la partie amont).

6.4 Réflexions sur la modularisation de la connaissance

6.4.1 Pourquoi modulariser la connaissance ?

L'une des difficultés rencontrées lors de la présentation des ontologies réside en la mythification du mot. On pourrait reprendre à propos des ontologies l'expression de Harold and Means (2002) au sujet du XML : *certains attendent tout de XML, y compris de laver le chien de la famille*. De la même manière, les ontologies sont souvent vues comme une solution universelle, une sorte de vague concept informatique magique qu'il suffirait de superposer à l'image pour qu'elle parle d'elle-même.

Mais l'image ne parle jamais d'elle-même. Que ce soit par le biais des ontologies ou bien par des approches classiques, l'interprétation d'une image satellitaire est toujours le fruit d'une connaissance experte (informations principale et extérieure de la FIGURE 2.6, page 43) et éventuellement de traitements d'image. Mais dans les deux cas, cela ne se fait pas tout seul et la machine n'invente rien.

Lorsqu'un expert propose un algorithme d'analyse d'image satellitaire (comme par exemple Baraldi et al. (2006)), celui-ci mélange les deux aspects que sont :

1. la structure algorithmique (par exemple, un arbre de décision) ;
2. la connaissance experte (par exemple les conditions de passage dans une branche ou l'autre de l'arbre).

Si la connaissance experte en télédétection est séparée de l'algorithme, celui-ci gagne en généricité. C'est le cas des algorithmes de raisonnement qui ne contiennent qu'une connaissance experte en logique de description, indépendante de tout autre domaine d'expertise qui utiliserait cette logique.

Si les ontologies de représentation de la connaissance experte sont contenues dans des fichiers, on obtient alors un traitement qui reçoit une connaissance formalisée très modulaire et qui peut être remplacée comme un simple paramètre. Le prix à payer pour cette modularité est une complexité génératrice de temps de calculs plus importants.

6.4.2 Comment modulariser la connaissance ?

C'est donc cette capacité des ontologies à capitaliser une connaissance logique qui a été utilisée, tout en restant pragmatique. En effet, une même connaissance experte peut être représentée à différents niveaux. Reprenons l'exemple de la connaissance décrite par Baraldi et al. (2006) sous la forme d'un arbre de décision.

À un premier niveau, l'arbre attribue simplement une classe (généralement d'occupation du sol) à chaque pixel de l'image. Mais une classe est déterminée au moyen de calculs, d'opérations logiques et d'appels à des procédures. Ces procédures contiennent elles-mêmes des calculs, des opérations logiques et des appels à des procédures... jusqu'à atteindre les valeurs des pixels.

Trois cas de figure de représentation de la connaissance sont envisageables. Chacune a des conséquences sur la connaissance à laisser incluse dans le traitement :

- La connaissance est représentée à un haut niveau d'abstraction. Les classes ne sont pas détaillées par des règles dans l'ontologie et la classification est effectuée par un algorithme dédié.
- La connaissance est représentée à un bas niveau d'abstraction. Les classes sont détaillées par des règles uniquement à partir des valeurs des pixels de l'image. Cela nécessite que les classes soient décrites dans l'ontologie par des règles et des calculs à l'aide de représentations mathématiques complexes comme les fonctions agrégatives (moyenne, maximum...). Cela suppose également que les raisonneurs implémentent ce genre d'opération. Or, d'après Baader and Sattler (1998), l'intégration de fonctions agrégatives rend rapidement le langage $\mathcal{ALC}_{\mathcal{D}\text{agg}}$[1] indécidable quant aux problèmes de satisfaisabilité et de subsomption.
- La connaissance est représentée à un niveau d'abstraction intermédiaire en utilisant uniquement des opérateurs logiques supportés par les standards et les raisonneurs éprouvés. Les opérateurs combinent non pas seulement les valeurs d'origine des pixels, mais également des valeurs produites par des traitements d'images (calculs d'indices, de textures...). Le traitement d'image garde la main sur les opérations arithmétiques, qui se trouvent relever d'une partie de la connaissance relativement objective (le calcul d'indices). Les opérations logiques sont déléguées aux raisonneurs.

Nous avons vu que la troisième solution a été privilégiée, de manière à se rapprocher le plus possible de l'image tout en gardant une approche pragmatique qui n'ignore ni les considérations techniques, ni les difficultés théoriques.

Rapprocher la connaissance structurée le plus possible de l'image est une des exigences indirectes de la thèse, puisque le but est d'établir un pont entre l'image et l'expertise. Ceci ne doit pas éclipser les considérations techniques, car le projet CARTAM-SAT vise des objectifs de recherche en lien étroit avec la conception pratique d'une chaîne de traitements intégré de l'image, depuis sa production en pied d'antenne jusqu'à sa diffusion.

6.5 Réflexions sur la représentation des relations spatiales

En évoquant la segmentation d'image, quelques considérations sur les limites des objets nous ont déjà mené à l'évocation des relations qu'ils peuvent entretenir entre eux. Ce point faisant partie des développements explorés dans la mise en œuvre, il importe de s'y attarder quelque peu.

6.5.1 Représentation des relations topologiques

Le choix de la représentation de ces relations s'est porté sur RCC8, car elle est présente dans les standards de l'OGC. Elle est également décrite dans l'ontologie GeoSPARQL. D'autre part, argument technique, *Orfeo Toolbox* implémente le calcul des relations RCC8 entre les régions de deux images segmentées. En fournissant deux fois la même image on obtient les relations topologiques entre les objets d'une seule image.

Nous avons observé en particulier les caractéristiques de la relation *Externally Connected*.

[1] L'ajout de l'indice $_{\text{agg}}$ à la notation du langage indique qu'il intègre des fonctions agrégatives.

6.5.2 Interprétation de l'adjacence dans une image matricielle avec RCC8.

Mais l'implémentation de RCC8 pour des images matricielles peut causer des surprises du fait de la définition de la frontière grâce à une dilatation par morphologie mathématique.

Quelles conditions deux segments S_1 et S_2 d'une image, doivent-ils remplir pour être considérés comme en relation *Externally Connected* ?

- Ils doivent chacun inclure des pixels n'appartenant pas à l'autre ($S_1 \cap S_2^{\complement} \neq \emptyset \wedge S_1^{\complement} \cap S_2 \neq \emptyset$).
- Ils ne doivent partager aucun pixel en commun ($S_1 \cap S_2 = \emptyset$).
- Leurs frontières doivent partager au moins un pixel.

D'après la définition 39 (page 53), les deux premières conditions sont réunies d'office sur une image segmentée. Reste la troisième. Elle repose sur la définition de frontières des segments. Or celle-ci n'est pas intuitive car elle fait appel à l'opération de dilatation en morphologie mathématique. La frontière des segments est constituée de la marge d'un pixel d'épaisseur qui entoure le segment[1]. Cette définition peut conduire à considérer comme *Externally Connected*, deux segments qui ne se touchent pas (FIGURE 6.8).

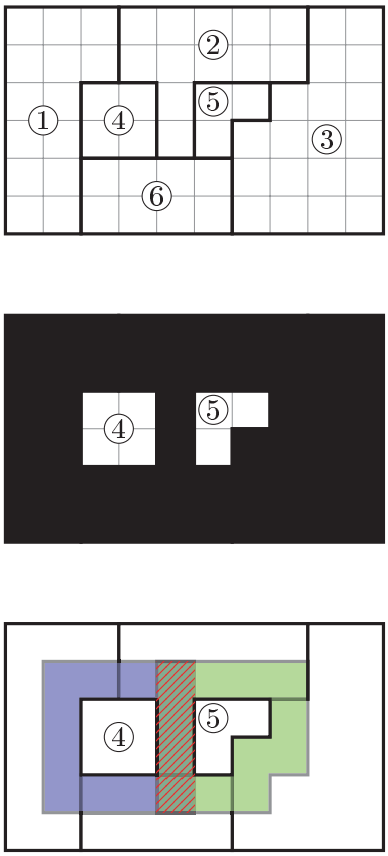


FIGURE 6.8 – Application de la relation RCC8 *Externally Connected* aux images matricielles. En considérant la partition de l'image matricielle proposée en haut, on désire savoir si les segments 4 et 5 (qui ne sont pas adjacent au sens commun de ce mot) sont liés par la relation RCC8 *Externally Connected*. Il faut pour cela les isoler comme (présenté au milieu), avant d'appliquer une dilatation d'un pixel correspondant à la définition de leurs frontières (représentation du bas). On constate alors que l'intersection des frontières n'est pas vide, ce qui implique que les segments sont *Externally Connected*.

[1] http://orfeo-toolbox.org/doxygen-current/otbImageToImageRCC8Calculator_8txx_source.html[accessed : 09-07-2013]

```
<?xml version="1.0"?>

<!DOCTYPE rdf:RDF [
    <!ENTITY rcc8 "http://www.orfeo-toolbox.org/otb/RCC8#" >
    <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
    <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
    <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]>

<rdf:RDF xmlns="http://www.w3.org/2002/07/owl#"
     xml:base="http://www.orfeo-toolbox.org/otb/RCC8"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
     xmlns:rcc8="http://www.orfeo-toolbox.org/otb/RCC8#"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
    <Ontology rdf:about="http://www.orfeo-toolbox.org/otb/RCC8#"/>

    <ObjectProperty rdf:about="&rcc8;disconnected"/>

    <ObjectProperty rdf:about="&rcc8;equal"/>

    <ObjectProperty rdf:about="&rcc8;externallyConnected">
        <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#
         SymmetricProperty"/>
    </ObjectProperty>

    <ObjectProperty rdf:about="&rcc8;nonTangentialProperPart"/>

    <ObjectProperty rdf:about="&rcc8;nonTangentialProperPartInverse">
        <inverseOf rdf:resource="&rcc8;nonTangentialProperPart"/>
    </ObjectProperty>

    <ObjectProperty rdf:about="&rcc8;partiallyOverlapping"/>

    <ObjectProperty rdf:about="&rcc8;tangentialProperPart"/>

    <ObjectProperty rdf:about="&rcc8;tangentialProperPartInverse">
        <inverseOf rdf:resource="&rcc8;tangentialProperPart"/>
    </ObjectProperty>
</rdf:RDF>
```

CODE 6.1 – Implémentation des relations rcc8 calquée sur la structure proposée par GeoSPARQL.

Hors de toute considération temporelle, *Externally Connected* est sans doute la relation RCC8 la plus intéressante à considérer entre les segments d'*une* image. Toute superposition totale ou partielle étant exclue entre les segments d'une même image, restent en effet à considérer *Externally Connected* et *Disconnected*. Mais cette dernière étant la plus répandue, c'est *Externally Connected* qui est la plus caractéristique et discriminante. Sur des images matricielles, il faut toutefois considérer avec prudence son interprétation comme l'*adjacence* entre segments, bien que cette tentation soit forte.

Afin d'éviter de surcharger le raisonneur par des concepts inutiles, nous avons repris de l'ontologie GeoSPARQL uniquement les propriétés représentant les relations spatiales (CODE 6.1).

7

Déclinaisons du principe et implémentations

Guider l'interprétation de l'image satellitaire par raisonnement sur les logiques de description est un principe général qui peut se décliner en pratique de nombreuses manières. Nous en proposons quatre variations techniques dans ce chapitre.

Sommaire

7.1 Application associée à l'analyse orientée objet **126**
7.1.1 Analyse orientée objet avec *Orfeo Toolbox* 126
7.1.2 Description de l'image segmentée 127
7.1.3 Exploitation de la description de l'image segmentée 128
7.1.4 Proposition d'évaluation 131
7.1.5 Conclusion de la validation de l'approche 134
7.2 Intégration des relations spatiales **135**
7.2.1 Vers la définition de nouveaux concepts 135
7.2.2 Proposition d'évaluation 136
7.2.3 Limites de l'utilisation de RCC8 138
7.2.4 Potentiel de la parallélisation 140
7.2.5 Influence de la segmentation et des caractéristiques de l'image 140
7.3 Adaptation à la description des pixels **141**
7.3.1 Une déclinaison *a priori* difficile 141
7.3.2 Contourner la complexité des logiques de description par une hypothèses de travail 142
7.3.3 Vers une segmentation ontologique ? 143
7.4 Classification séparée des objets **145**
7.4.1 Comparaison des classifications globale et séparée 145
7.4.2 Produits du raisonnement : du masque sémantique à la carte interprétée 145
7.5 Réflexions **148**
7.5.1 Valeur ajoutée de l'approche 148
7.5.2 Les enjeux du passage à l'échelle 148
7.5.3 Vers une approche ontologique hybride pixels/objets ? 148
7.6 L'hypothèse du monde ouvert **149**
7.6.1 Comment définir les concepts ? 149
7.6.2 Comment décrire un fait ? 152
7.6.3 Comment regarder les résultats ? 152

Les quatre premières sections sont consacrées à la présentation d'une déclinaison pratique du principe d'interprétation par raisonnement automatique. Leur succession est à la fois logique et globalement chronologique.

1. La première consiste à se placer à la suite d'une analyse orientée objet et à raisonner sur la description des segments. Elle nécessite une segmentation qui réduit le nombre d'entités à décrire. Ce dernier point est important si l'on considère les questions de complexité des logiques de description.

2. La seconde déclinaison ajoute la considération des relations spatiales et conduit à se poser la question de leur calcul et de leur signification. Elle oblige à réviser la manière de confronter les résultats obtenus à ceux produits par d'autres moyens.

3. La troisième propose une hypothèse supplémentaire et une astuce technique afin d'appliquer le raisonnement à la description des pixels sans ruiner les performances.

4. La dernière proposition d'application revient à une approche objet en y transposant l'astuce utilisée au niveau des pixels.

Ces quelques explorations techniques amènent des réflexions sur ce qui pourrait être une approche tenant compte des avantages et des inconvénients de chacun des aspects explorés : pixel, objet, spatial. Elles posent également la question de la manière dont les concepts doivent être définis et les résultats interprétés.

7.1 Application associée à l'analyse orientée objet

La première application n'avait d'autre but que de vérifier la faisabilité technique de l'approche. Est-elle fondée ? Quelles sont les questions qui se présentent ? Quels problèmes techniques se posent-ils ? Dans une certaine mesure, il s'agit d'une simple validation expérimentale.

L'application se place dans le cadre d'une analyse d'image orientée objet.

Hypothèse 6 (Hypothèse de description des segments)**.** Il s'agit de décrire les caractéristiques des segments d'image de manière *quantitative* en se basant sur une conceptualisation de référence.

La description de tous les segments de l'image alimente une même base de connaissance.

Hypothèse 7 (Hypothèse d'interprétation par raisonnement)**.** L'interprétation est réalisée par raisonnement sur une base de connaissances contenant la conceptualisation de référence globale, la connaissance contextuelle et la description de l'ensemble des segments de l'image.

Détaillons maintenant dans ses aspects techniques la manière dont l'image est traitée puis décrite, avant que sa description ne soit elle-même exploitée par raisonnement.

7.1.1 Analyse orientée objet avec *Orfeo Toolbox*

Orfeo Toolbox[1] (OTB) est une bibliothèque logicielle libre[2] de traitements d'images de télédétection. Elle s'appuie sur *Insight Segmentation and Registration Toolkit*[3] (ITK), bibliothèque de traitement d'imagerie médicale également libre[4].

[1] http://www.orfeo-toolbox.org/otb/.
[2] *Orfeo Toolbox* est distribuée sous licence CeCILL (http://orfeo-toolbox.org/otb/).
[3] http://www.itk.org/.
[4] Depuis la version 4.0, *Insight Toolkit* est distribuée sous licence Apache 2.0 (http://www.itk.org/ITK/project/license.html).

Entre autres, ITK propose des classes pour la lecture et l'écriture de fichiers d'images indépendamment du format de fichier. OTB étend ITK. Certaines classes sont spécialisées et d'autres sont ajoutées pour s'adapter aux spécificités de la télédétection. Les traitements d'image basiques tels que les calculs de textures sont directement fournis par ITK. Le code source est implémenté en C++. Ce choix d'un langage répandu permet d'espérer une combinaison de l'interface de programmation d'OTB avec d'autres bibliothèques dans le même langage.

Le développement d'OTB est encouragé par le Centre National d'Études Spatiales (CNES) et son utilisation est recommandée dans le cadre du projet CARTAM-SAT.

La partie en amont de l'approche qui consiste à effectuer une analyse orientée objet est confiée à OTB (FIGURE 7.1).

Une première étape suivant la lecture de l'image consiste en une simple analyse orientée pixel : l'extraction des bandes spectrales, le calcul des indices spectraux et texturaux auxquels on s'intéresse dans l'image.

Puis, l'image est segmentée de manière à obtenir une *bonne segmentation* telle que nous l'avons définie. L'image de segmentation qui en résulte, contient des segments d'image à partir desquels il est possible de calculer des caractéristiques de forme rangées dans une structure de données. Ces caractéristiques sont accessibles pour chaque segment à partir de son identifiant (La FIGURE 7.2 représente ces structures dans la partie gauche.).

Ensuite, d'autres classes d'OTB permettent de combiner à l'image de segmentation chacune des caractéristiques calculées sur l'ensemble de l'image lors de l'étape d'analyse orientée pixel. Ces caractéristiques sont accessibles pour chaque segment, dans d'autres structures de données, grâce à son identifiant (FIGURE 7.1).

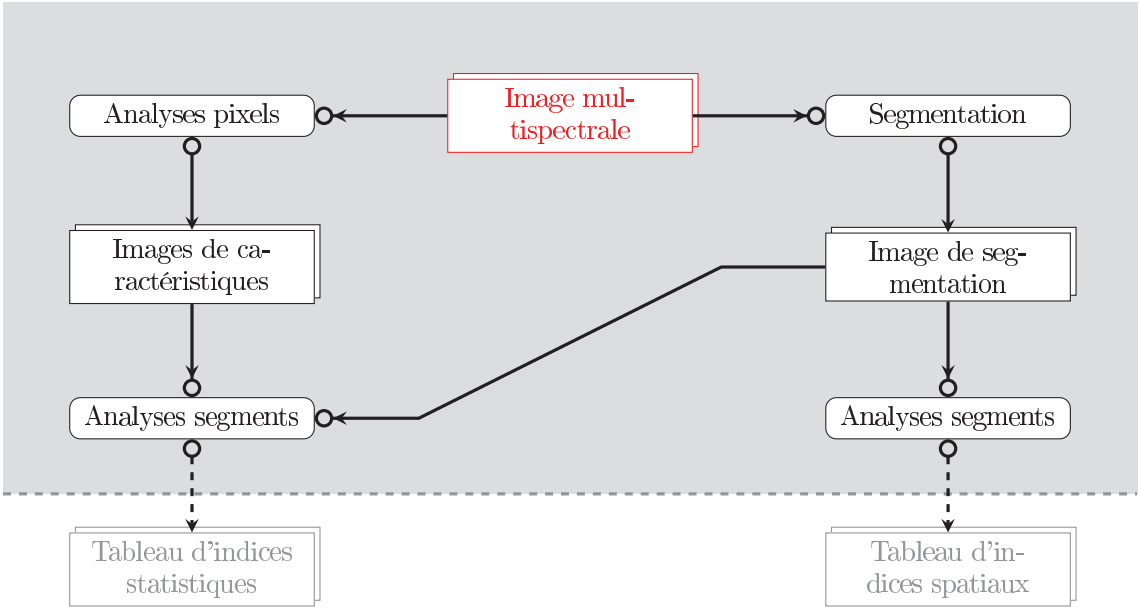


FIGURE 7.1 – Rappel de la partie amont (analyse de l'image) des chaînes de traitements exposées en FIGURE 6.5, FIGURE 6.6 et FIGURE 6.7.

7.1.2 Description de l'image segmentée

Chaque tableau de caractéristiques est passé en revue de manière à décrire les principales caractéristiques calculées pour tous les segments de l'image. Puis chaque structure décrite est rattachée dans la description à un concept de la conceptualisation de référence (FIGURE 7.2).

Mais quelle forme donner à cette description ? La première idée serait de la rendre conforme aux standards actuels du génie ontologique. La bibliothèque *Soprano* permet de réaliser cette étape pour des descriptions réduites. C'est ce que nous avons qualifié plus haut de proposition *non-intégrée*.

Dans cette proposition *non-intégrée*, chaque description sous forme de triplet est insérée dans un modèle de connaissances de *Soprano*. Le passage en OWL au format RDF/XML est pris en charge par la bibliothèque. Il en résulte en sortie un fichier OWL descriptif des segments de l'image analysée en termes de caractéristiques de l'image.

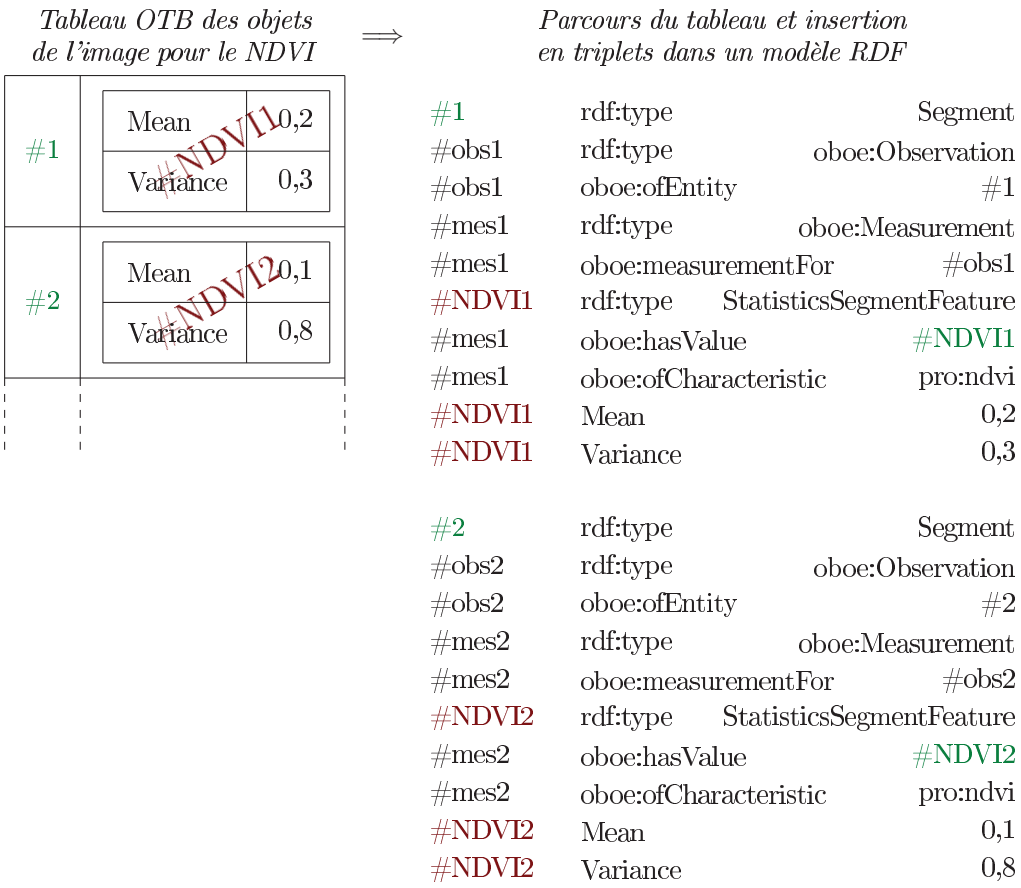


FIGURE 7.2 – Schéma simplifié de la description ontologique de l'image avec *Soprano* (version *non-intégrée*). La description des faits est similaire en ce qui concerne l'alimentation d'un raisonneur (version *intégrée*). Au lieu d'alimenter un modèle RDF, les triplets prennent la forme d'assertions qui alimentent la base de connaissances du raisonneur.

Le principe est exactement le même lorsque l'outil *Soprano* est remplacé par l'interface de programmation du raisonneur FaCT++ de manière à élaborer les propositions *intégrées*. Chaque entité de l'image référencée dans les structures issues du traitement par OTB est décrite et rejoint la base de connaissances qui alimente le raisonneur[1].

7.1.3 Exploitation de la description de l'image segmentée

L'exposé du principe général de l'approche fait ressortir la nécessité d'une connaissance experte pour l'obtention d'une interprétation sémantique de l'image par un raisonneur. En effet, l'interprétation n'est autre que la classification des instances représentant les entités de l'image dans des concepts dotés d'une sémantique de télédétection (FIGURE 7.3).

Il importe donc maintenant de définir ces concepts. De manière à vérifier la faisabilité de l'approche et à valider des résultats concrets, nous proposons de nous appuyer sur trois concepts principaux représentant les segments de végétation :

[1]Le vocabulaire OWL, apparent lors de la manipulation de *Soprano*, disparaît lors de l'utilisation directe de l'interface de programmation d'un raisonneur pour laisser place à des notions de logique de description.

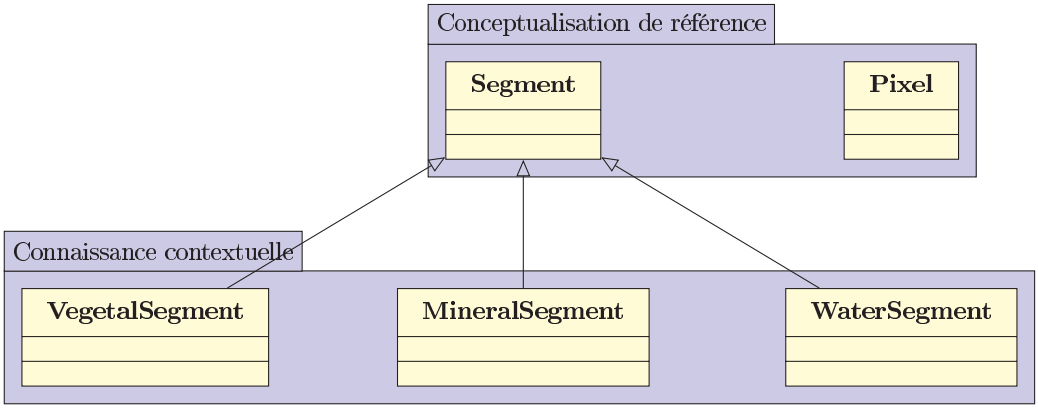


FIGURE 7.3 – Le segment d'image et ses interprétations sémantiques.

$$\begin{aligned} VegetatedSegment \doteq & Segment \\ & \sqcap \exists observedBy.(Observation \\ & \sqcap \exists hasMeasurement.(Measurement \\ & \sqcap \exists hasValue.(StatisticsSegmentFeature \\ & \sqcap \exists Mean.\{double \geq 0,3\} \\ & \sqcap \exists Mean.\{double \leq 0,8\}) \\ & \sqcap \exists ofCharacteristic.\{ndvi\}))) \end{aligned}$$

...les segments d'eau :

$$\begin{aligned} WaterSegment \doteq & Segment \\ & \sqcap \exists observedBy.(Observation \\ & \sqcap \exists hasMeasurement.(Measurement \\ & \sqcap \exists hasValue.(StatisticsSegmentFeature \\ & \sqcap \exists Mean.\{double \geq 0,0\}) \\ & \sqcap \exists ofCharacteristic.\{ndwi2\}))) \end{aligned}$$

...et les segments de surfaces minérales ou construites :

$$\begin{aligned} MineralSegment \doteq & Segment \\ & \sqcap \exists observedBy.(Observation \\ & \sqcap \exists hasMeasurement.(Measurement \\ & \sqcap \exists hasValue.(StatisticsSegmentFeature \\ & \sqcap \exists Mean.\{double \geq 0,0\}) \\ & \sqcap \exists ofCharacteristic.\{ndbi\}))) \end{aligned}$$

Ces définitions exprimées ici en utilisant le formalisme des logiques de description, ont également été implémentées en utilisant le vocabulaire OWL dans un fichier RDF/XML (CODE 7.1).

Une fois le raisonneur alimenté avec la connaissance nécessaire issue de l'image et de l'expertise, il ne reste plus qu'à raisonner sur la base de connaissances afin de la réaliser. De cette manière, les instances qui sont des entités de l'image se voient attribuer les concepts qui leur sont les plus spécifiques, c'est-à dire dans la mesure du possible, les sémantiques de télédétection définies par l'expert comme autant de spécialisations du concept de *Segment*.

```
<Class rdf:about="&teledetection;VegetatedSegment">
 <equivalentClass>
  <Class>
   <intersectionOf rdf:parseType="Collection">
    <rdf:Description rdf:about="&img;Segment"/>
    <Restriction>
     <onProperty rdf:resource="&oboe-core;observedBy"/>
     <someValuesFrom>
      <Class>
       <intersectionOf rdf:parseType="Collection">
        <rdf:Description rdf:about="&oboe-core;Observation"/>
        <Restriction>
         <onProperty rdf:resource="&oboe-core;hasMeasurement"/>
         <someValuesFrom>
          <Class>
           <intersectionOf rdf:parseType="Collection">
            <rdf:Description rdf:about="&oboe-core;Measurement"/>
            <Restriction>
             <onProperty rdf:resource="&oboe-core;hasValue"/>
             <someValuesFrom>
              <Class>
               <intersectionOf rdf:parseType="Collection">
                <rdf:Description rdf:about="&img;StatisticsSegmentFeature"/>
                <Restriction>
                 <onProperty rdf:resource="&img;Mean"/>
                 <someValuesFrom>
                  <rdfs:Datatype>
                   <onDatatype rdf:resource="&xsd;double"/>
                   <withRestrictions rdf:parseType="Collection">
                    <rdf:Description>
                     <xsd:minInclusive rdf:datatype="&xsd;double">0.3
                     </xsd:minInclusive></rdf:Description>
                   </withRestrictions></rdfs:Datatype>
                 </someValuesFrom></Restriction>
                <Restriction>
                 <onProperty rdf:resource="&img;Mean"/>
                 <someValuesFrom>
                  <rdfs:Datatype>
                   <onDatatype rdf:resource="&xsd;double"/>
                   <withRestrictions rdf:parseType="Collection">
                    <rdf:Description>
                     <xsd:maxInclusive rdf:datatype="&xsd;double">0.8
                     </xsd:maxInclusive></rdf:Description>
                   </withRestrictions></rdfs:Datatype>
                 </someValuesFrom></Restriction>
               </intersectionOf>
              </Class>
             </someValuesFrom></Restriction>
            <Restriction>
            <onProperty rdf:resource="&oboe-core;ofCharacteristic"/>
            <hasValue rdf:resource="&pro;ndvi"/>
            </Restriction>
           </intersectionOf>
          </Class>
         </someValuesFrom>
        </Restriction>
       </intersectionOf>
      </Class>
     </someValuesFrom>
    </Restriction>
   </intersectionOf>
  </Class>
 </equivalentClass>
 <rdfs:subClassOf rdf:resource="&img;Segment"/>
</Class>
```

CODE 7.1 – La représentation du concept de segment de végétation (*VegetatedSegment*) dans le fichier RDF/XML contenant la connaissance experte.

Le résultat montre que cette approche donne des résultats cohérents que l'on peut l'observer sous diverses formes.

L'exemple proposé en FIGURE 7.4, illustre l'aboutissement de la proposition *non-intégrée* : les résultats sont accessibles à travers l'éditeur d'ontologies *Protégé*. Dans cet essai simplifié, le seul segment correspondant à la définition du concept de segment végétal est classifié[1] comme tel.

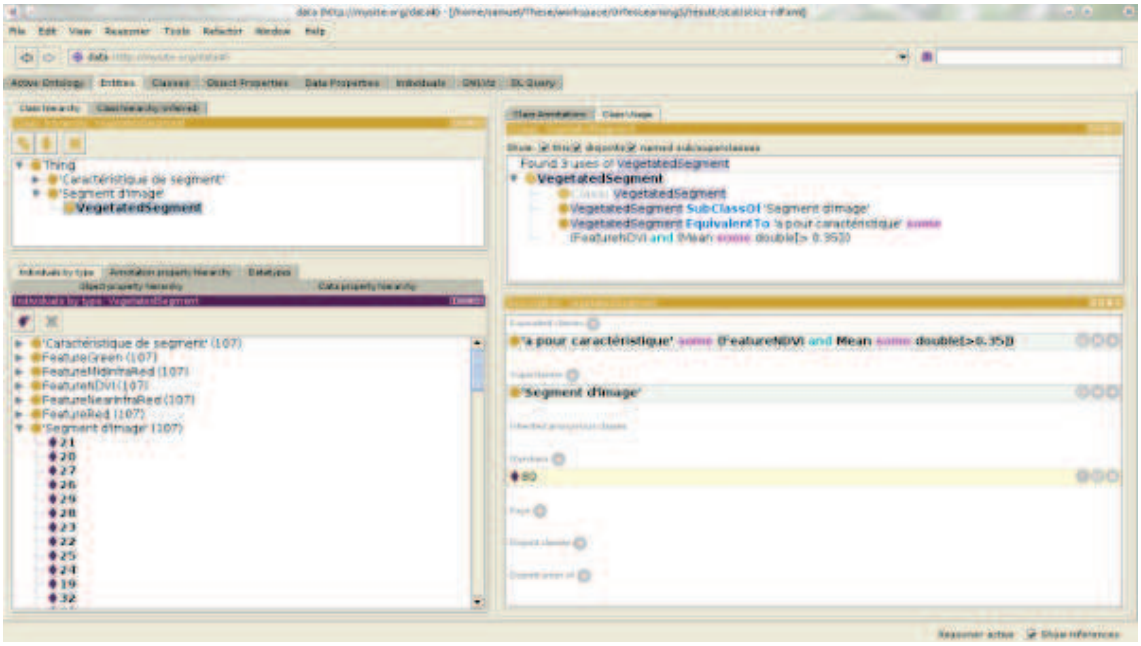


FIGURE 7.4 – L'éditeur d'ontologies *Protégé*. Le cadre inférieur droit montre un résultat inféré : l'individu représentant le segment d'image #80 a été classifié dans le concept de segment végétal, sélectionné dans le cadre supérieur gauche.

Mais grâce à l'intégration du raisonnement, il est possible de visualiser les résultats sous forme d'images numériques. La FIGURE 7.5 montre par exemple sous la forme d'un « masque » quelles sont les zones de l'image qui correspondent à des segments de végétation.

7.1.4 Proposition d'évaluation

Évaluer ce résultat est une tâche délicate dont la pertinence se discute. Nous préférons voir le résultat comme une interprétation qui en vaut bien une autre, plutôt que de le comparer à des méthodes qui ne sont pas forcément des références indiscutables.

Une autre difficulté dans l'évaluation de chaînes de traitements est de bien comprendre *ce qui est réellement évalué*. Dans le cas présent, de nombreuses étapes sont susceptibles de l'être : la segmentation, le raisonnement, l'élicitation de la connaissance experte... La modification de chacune peut influer sur le résultat.

C'est à une évaluation générale de la méthode que nous souhaitons procéder et non à l'évaluation d'une étape particulière supposée satisfaisante par hypothèse. La segmentation est supposée *bonne*, les définitions expertes sont *a priori* correctes et le raisonneur n'a aucune raison de donner des résultats absurdes. Mais que dire de la combinaison de tous ces éléments et comment les résultats s'accordent-ils aux produits d'autres *méthodes* ?

Pour cela, nous avons effectué quelques comparaisons statistiques entre les masques thématiques obtenus par raisonnement sur les segments et des masques thématiques obtenus par simple seuillage

[1]Les difficultés de vocabulaire inhérentes aux applications pluridisciplinaires nous rejoignent ici. En ce qui concerne la représentation des connaissances, l'individu représentant le segment est *réalisé* dans le concept de *VegetatedSegment*. Mais en ce qui concerne la télédétection, le segment est *classifié* comme segment de végétation. Classification, réalisation : il se peut donc que ces mots soient assimilés par la suite, en ce qui concerne l'interprétation sémantique des entités de l'image par raisonnement.

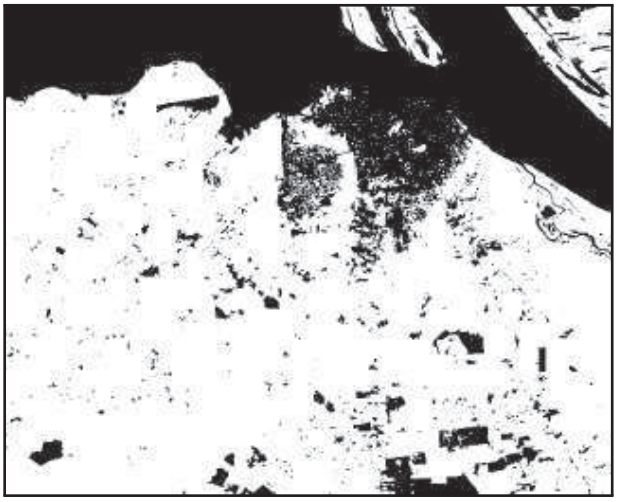


FIGURE 7.5 – Visualisation, en blanc, des segments de l'extrait (825 × 666 pixels) de l'image Landsat 5, dotés d'une sémantique végétale. Les segments en noir n'ont pas été classifiés dans cette sémantique.

de pixels. Pour chaque comparaison, ce sont les mêmes seuils et les mêmes indices qui ont été utilisés. Cette évaluation inclut donc la combinaison de la segmentation, du raisonnement et de la traduction des règles de l'expertise avec les logiques descriptives[1], puisque toutes ces étapes diffèrent entre les deux méthodes comparées. Par contre, elle n'évalue pas la pertinence des seuils fixés aux indices, puisqu'ils sont identiques des deux côtés.

Lorsque l'on souhaite comparer les résultats de deux méthodes, il est intéressant de construire des *matrices de confusion*. Ces tableaux recensent la part d'éléments sur lesquels les deux méthodes s'accordent ou se contredisent. Les matrices de confusion sont particulièrement indiquées lorsque les résultats prennent la forment d'éléments classés *positifs* ou *négatifs*. C'est presque le cas en ce qui nous concerne : dans un masque, une entité de l'image est classée, ou pas, comme représentative d'une sémantique de télédétection donnée. Cependant, deux précisions doivent être apportées avant de présenter ces matrices.

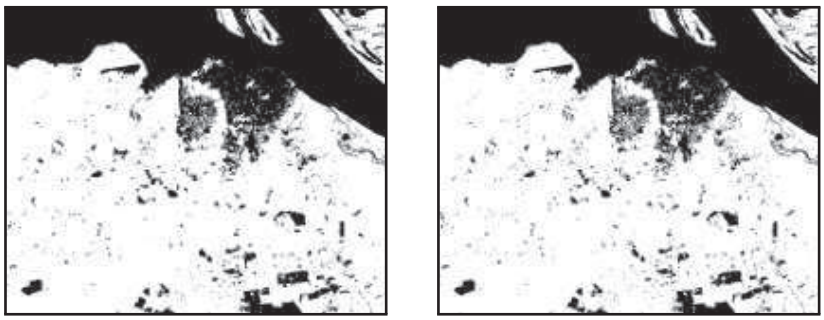
- Tout d'abord, l'une des méthodes classifie des segments et l'autre classifie des pixels. Si l'on reporte les contours des segments sur le résultat de la méthode basée sur les pixels, il est fort probable que certains d'entre eux seront composés de pixels classifiés différemment. Il semble donc pertinent de comparer entre eux les plus petits éléments : les pixels.

- Les matrices de confusion comparent habituellement des éléments classés par deux méthodes comme des exemples positifs ou négatifs. C'est le cas pour la méthode de seuillage par pixel. Mais nous verrons que l'hypothèse de monde ouvert ne permet pas d'en dire autant du résultat obtenu par raisonnement. Pour cette raison, un point d'interrogation désigne la colonne qui devrait correspondre aux exemples négatifs produits par raisonnement.

Les FIGURE 7.6 à FIGURE 7.8 illustrent notre démarche de « validation ». Le fait que les grandes sémantiques ressortent bien est confirmé par la technique élémentaire du seuillage de pixels. Cet accord est confirmé par des valeurs statistiques, comme le proposent les TABLE 7.1 à TABLE 7.3 avec des indices élaborés à partir des matrices de confusion.

La *précision totale* indique le taux d'éléments sur la classification desquels les deux méthodes s'accordent. L'indice κ dit « de Cohen » (Cohen et al., 1960) donne une évaluation de cet accord par rapport à ce qui pourrait être un accord aléatoire. Entre 0 et 1, l'accord de deux méthodes augmente lorsque le κ s'approche de 1. Un κ inférieur à 0 indique un désaccord des méthodes.

[1]La connaissance est la même pour les deux méthodes, mais elle fait l'objet d'une élicitation ontologique d'un côté tandis qu'elle conserve la forme de règles traditionnelles de l'autre.



(A) (B)

FIGURE 7.6 – Sémantique de végétation, obtenue par raisonnement sur des objets (A) et par seuillage de pixels (B). Les entités végétales sont en blanc.

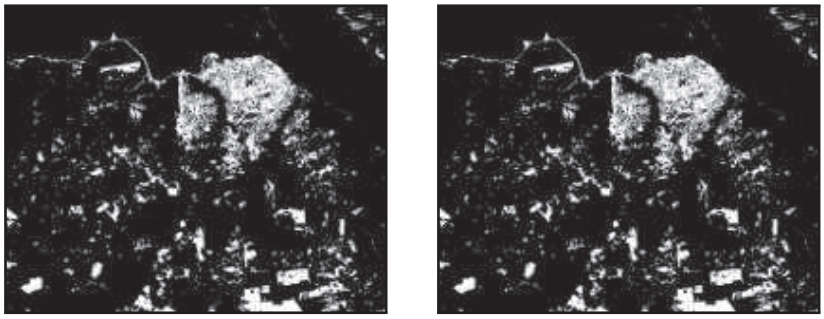
Matrice de confusion

seuillage \ raisonnement	?	végétal
non végétal	154089	2577
végétal	5885	386899

précision totale $= 0,985$

$\kappa = 0,962$

TABLE 7.1 – Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et l'approche par seuillage de pixels (comparaison pixel à pixel des images en FIGURE 7.6).



(A) (B)

FIGURE 7.7 – Sémantique de surfaces minérales ou construites, obtenue par raisonnement sur des objets (A) et par seuillage de pixels (B). Les entités minérales ou construites sont en blanc.

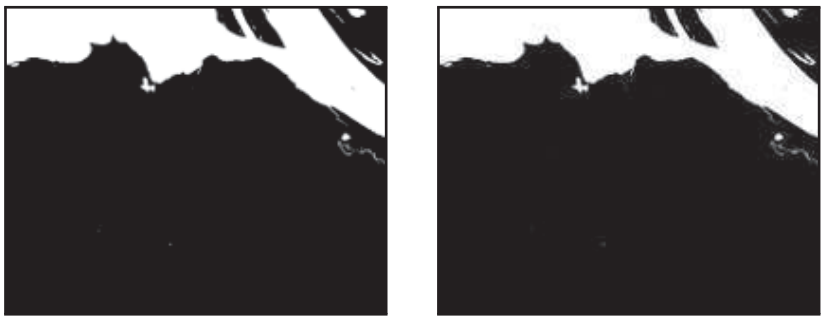
Matrice de confusion

seuillage \ raisonnement	?	minéral
non minéral	488019	134
minéral	170	61127

précision totale $= 0,999$

$\kappa = 0,997$

TABLE 7.2 – Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et l'approche par seuillage de pixels (comparaison pixel à pixel des images en FIGURE 7.7).



(A) (B)

FIGURE 7.8 – Sémantique aquatique, obtenue par raisonnement sur des objets (A) et par seuillage de pixels (B). Les surfaces en eau sont en blanc.

Matrice de confusion

seuillage \ raisonnement	?	eau
non eau	448746	263
eau	413	100028

précision totale $= 0,999$

$\kappa = 0,996$

TABLE 7.3 – Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et l'approche par seuillage de pixels (comparaison pixel à pixel des images en FIGURE 7.8).

7.1.5 Conclusion de la validation de l'approche

Dans une première approche en aval d'une analyse d'image orientée objet, le principe a pu être validé. On peut le rapprocher de la proposition de Falomir et al. (2011), à deux différences notables. Concernant le champ d'application, tout d'abord, car la télédétection est un domaine très particulier avec des données volumineuses et une connaissance spécifique. Concernant la manière de décrire l'image, ensuite, car la description n'utilise pas de modèle qualitatif et reste quantitative, au plus proche de l'image.

Résultat 1 (Raisonnement sur une base de connaissances contenant une description complète de l'image suite à une analyse orientée objet)**.** L'implémentation expérimentale de validation de l'approche permet de conclure que :

1. Il est possible de donner une description sémantique d'une image segmentée en s'appuyant sur une conceptualisation de référence de l'image satellitaire.
2. Une conceptualisation de référence permet également de produire une connaissance contextuelle représentant une expertise humaine formalisée.
3. Il est possible de partager la description de l'image en s'appuyant sur les standards du Web sémantique. Ce partage s'entend principalement au sens de l'interopérabilité : transmettre la connaissance à un logiciel capable de l'exploiter.
4. En s'appuyant sur la conceptualisation de référence, la connaissance contextuelle et les faits de l'image, un raisonneur est capable de rattacher un individu représentant une entité de l'image à un concept doté d'une sémantique spécialisée en télédétection.

7.2 Intégration des relations spatiales

Alors que le premier développement de l'approche consiste à construire une base de connaissances contenant une description des entités de l'image basée sur leurs caractéristiques intrinsèques, il s'agit maintenant de la compléter par une information relative aux relations spatiales que les segments entretiennent entre eux. Nous nous intéressons en particulier aux relations spatiales topologiques RCC8 implémentées dans l'outil de traitement d'images *Orfeo Toolbox* que nous utilisons.

7.2.1 Vers la définition de nouveaux concepts

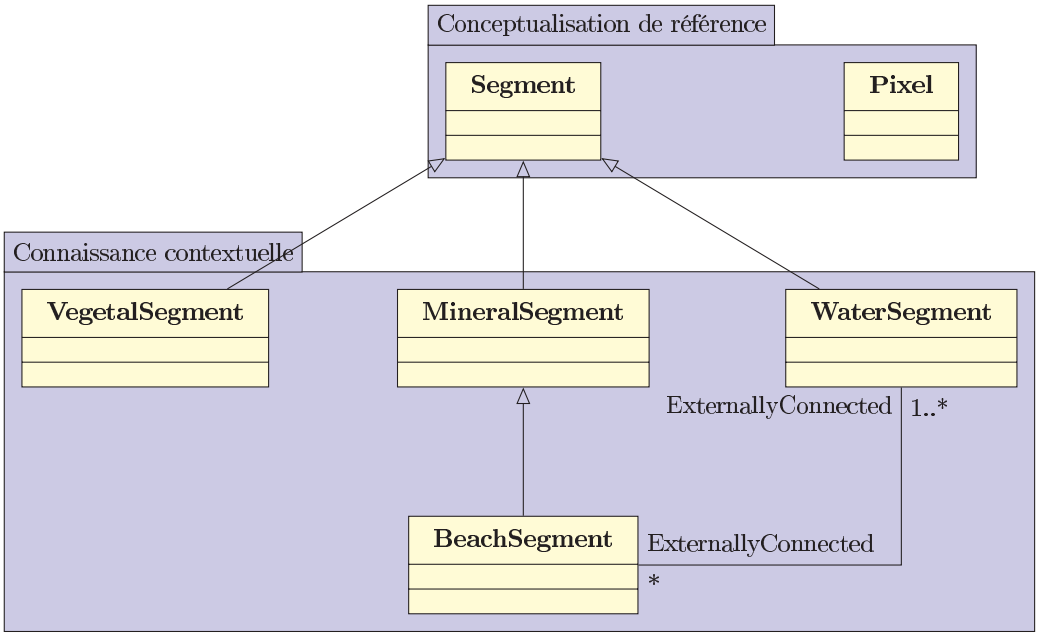


FIGURE 7.9 – Définition d'une sémantique de segment minéral *à proximité* d'un segment d'eau.

La prise en compte des relations spatiales permet de définir de nouveaux concepts prenant en compte cette nouvelle information extrinsèque aux segments (FIGURE 7.9). Nous avons pris comme exemple illustratif le concept de segment de *plage* ou de *rive* en le définissant comme un segment minéral *Externally Connected* à un segment d'eau (Andrés et al., 2013; Pierkot et al., 2013). De même que pour les descriptions basées sur les caractéristiques intrinsèques, une description utilisant les relations spatiales peut-être exprimée en utilisant un formalisme de logiques de descriptions :

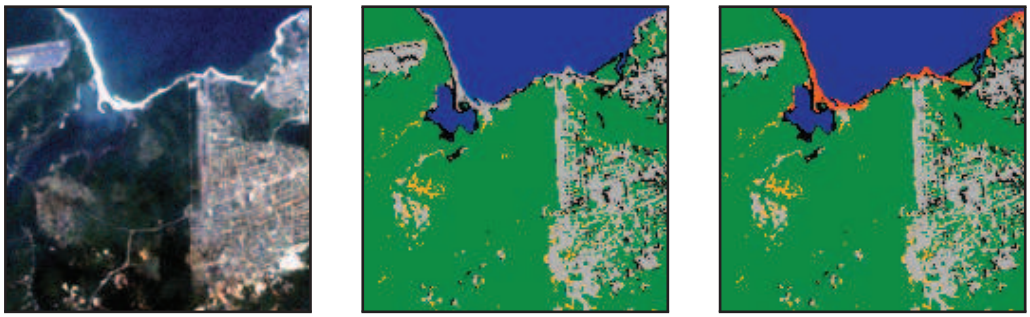
$$BeachSegment \doteq MineralSegment \sqcap \exists ExternallyConnected.WaterSegment$$

Il est également possible d'implémenter cette définition dans un fichier en utilisant les standards du web sémantique (CODE 7.2).

Dans une telle définition, les relations spatiales apparaissent comme un moyen de combiner des concepts discriminés par des caractéristiques intrinsèques, pour en construire de nouveaux. Cette création de nouveaux concepts, en offrant au télédétecteur la possibilité d'exprimer un vocabulaire plus détaillé, est synonyme d'enrichissement sémantique de l'interprétation. L'expert en télédétection peut ainsi affiner l'expression de sa connaissance. Bien entendu, en utilisant la même démarche que pour l'application précédente, les nouveaux concepts définis peuvent être détectés dans l'image (FIGURE 7.10).

```
<Class rdf:about="&teledetection;BeachSegment">
 <equivalentClass>
  <Class>
   <intersectionOf rdf:parseType="Collection">
    <rdf:Description rdf:about="&teledetection;MineralOrBuiltUpSegment"/>
    <Restriction>
     <onProperty rdf:resource="&rcc8;externallyConnected"/>
     <someValuesFrom rdf:resource="&teledetection;WaterSegment"/>
    </Restriction>
   </intersectionOf>
  </Class>
 </equivalentClass>
 <rdfs:subClassOf rdf:resource="&img;Segment"/>
</Class>
```

CODE 7.2 – La représentation du concept de segment de plage (*BeachSegment*) dans le fichier RDF/XML contenant la connaissance experte.



(A) (B) (C)

FIGURE 7.10 – Utilisation de la connaissance sur une relation spatiale dans un extrait d'image Landsat à proximité de Santarém (A). En (B), une interprétation automatique sans relations spatiales. Les segments d'eau sont représentés en bleu, les segments végétaux en vert et les segments minéraux ou construits en gris. Cette première interprétation permet de détecter des combinaisons de concepts de télédétection (végétation et minéral en jaune, eau et minéral en bleu clair) dont l'interprétation relève des experts du domaine. D'autres concepts peuvent être construits à l'aide de relations spatiales, par exemple pour la détection des rives (C). Les segments qui s'y rapportent sont représentés en orange.

7.2.2 Proposition d'évaluation

Cet exemple fournit l'occasion d'une nouvelle évaluation du principe. Pour un concept étroitement lié à la prise en compte de relations spatiales, cela n'aurait aucun sens de comparer les résultats obtenus à un seuillage de pixels[1]. Nous avons donc proposé (Pierkot et al., 2013) leur confrontation à des analyses réalisées manuellement ou semi-manuellement par des experts en télédétection (FIGURE 7.11).

En utilisant une nouvelle fois les matrices de confusion, le calcul de statistiques sur les masques obtenus fournit une évaluation quantitative. Cela mérite quelques commentaires.

L'indice κ de la première confrontation aux dires d'expert (TABLE 7.4) peut sembler décevant. Pourtant, la précision totale est bonne. Cela provient sans doute du concept lui-même qui désigne un ensemble d'objets relativement allongé. Cette configuration favorise par conséquent la classification approximative de pixels « à la main ». On peut ajouter que les données de l'analyse experte ont nécessité le passage d'un format vectoriel à au format matriciel. Cette étape peut également expliquer la perte de pixels en bordure de la représentation de la plage, ce qui peut éventuellement

[1]Cette méthode se limite aux valeurs des pixels, c'est-à-dire par définition à des caractéristiques intrinsèques.

(A) (B) (C)

FIGURE 7.11 – Sémantique de plage ou de rive obtenue par raisonnement sur des objets (A) et par l'expertise manuelle de deux experts en télédétection (B) et (C) (respectivement Mme Frédérique SEYLER et M. Jean-François FAURE). Les surfaces de rives sont en blanc.

Matrice de confusion

raisonnement / expertise 1	?	rive
non rive	39163	341
rive	172	324

précision totale $= 0,987$

$\kappa = 0,55$

TABLE 7.4 – Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et une interprétation réalisée manuellement par Mme Frédérique SEYLER.

réduire l'accord des résultats. On remarquera en tout cas que dans la première description experte, la rive est assez étroite, mais cela peut simplement relever de l'acception personnelle du concept par l'expert.

Matrice de confusion

raisonnement / expertise 2	?	rive
non rive	39089	187
rive	246	478

précision totale $= 0,989$

$\kappa = 0,68$

TABLE 7.5 – Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et une interprétation réalisée manuellement par M. Jean-François FAURE.

La seconde comparaison à une analyse experte (TABLE 7.5) produit de meilleurs indices de précision totale et de κ, mais le taux de faux négatifs est plus important que précédemment. Cela est vraisemblablement dû au fait de l'acception plus large du trait de côte par l'expert qui apparaît par conséquent plus épais sur le masque.

Enfin, une troisième comparaison est nécessaire. Elle concerne les deux analyses expertes (TABLE 7.6) et montre que, au sens de l'indice κ, l'analyse automatique est plus proche de la seconde analyse experte que les deux analyses expertes sont similaires entre elles. **Dans ce cas de figure et d'un point de vue statistique, cela confirme que l'interprétation ontologique par raisonnement automatique a toute sa place en télédétection.**

Matrice de confusion

expertise 2 \ expertise 1	non rive	rive
non rive	39171	105
rive	333	391

précision totale $= 0,989$

$\kappa = 0,64$

TABLE 7.6 – Matrice de confusion et statistiques dérivées comparant les interprétations réalisées manuellement par Mme Frédérique SEYLER et M. Jean-François FAURE.

7.2.3 Limites de l'utilisation de RCC8

7.2.3.1 Artefact sémantique

Il est tentant d'interpréter directement la relation *Externally Connected* comme indiquant l'adjacence de deux objets de l'image. C'est implicitement ainsi que le concept relatif à la détection des plages ou des rives a été défini. Nous avons vu plus haut que la relation *Externally Connected* ne pouvait pas être rigoureusement identifiée à l'adjacence (illustration pratique en FIGURE 7.12).

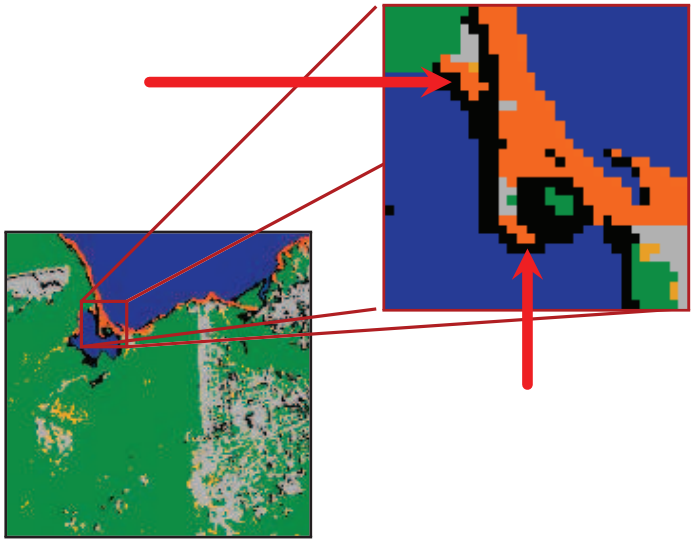


FIGURE 7.12 – Artefacts liés à l'implémentation du calcul des relations spatiales RCC8 dans les images matricielles. Deux segments de plage (en orange) ne sont adjacents à aucun segment d'eau.

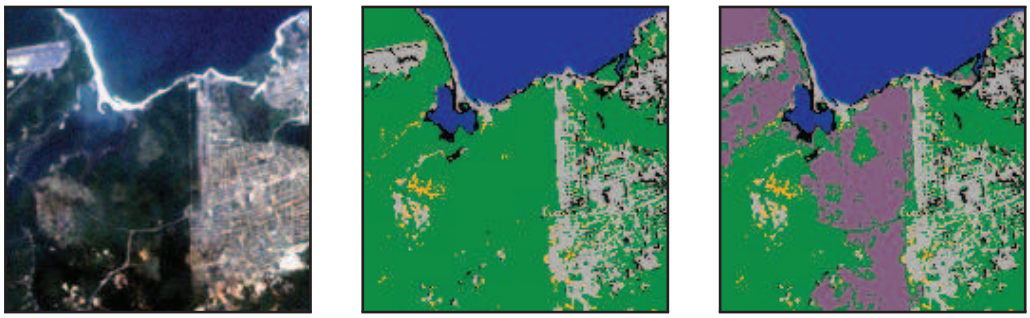
Enfin, il faut toujours se demander quels sont les concepts accessibles dans une image, d'une part en fonction de ses caractéristiques, et d'autre part en fonction du langage utilisé pour décrire son contenu et représenter la connaissance. Par exemple, sans les relations spatiales, certains concepts comme les bordures de fleuves, ne pourront pas être décrits, alors qu'ils seront toujours *visibles* sur l'image pour l'œil exercé.

Mais pour d'autres concepts, l'échec de leur détection dans une image dépassera la question du langage utilisé : ont-ils des caractéristiques (intrinsèques ou extrinsèques) qui leur sont propres parmi celles représentées par l'image ? Il peut alors être question d'une caractéristique spectrale qui nécessiterait une bande supplémentaire, ou d'une caractéristique spatiale (de forme ou de relation spatiale) qui nécessiterait d'augmenter la résolution spatiale. En d'autres termes : le contexte de validité de la connaissance est-il compatible avec les caractéristiques de l'image ? La connaissance est-elle *correctement* définie en considération des caractéristiques de l'image, de l'environnement représenté etc. ?

7.2.3.2 Utilisation non pertinente ou insuffisante

On peut illustrer l'importance du contexte de validité par un exemple d'utilisation non pertinente (ou au moins discutable) des relations spatiales. En milieu côtier, il est fréquent de s'intéresser à une végétation particulière se développant sur des bancs de sable qui se déposent sur la côte pour quelques années, puis disparaissent au gré des courants. Cette végétation qui supporte les intrusions d'eau salée causées par les marées, constitue un milieu intéressant à détecter sur des images satellites.

Le milieu ici n'est pas à proprement parler côtier, puisque Santarém se situe au confluent de l'Amazone et du Tapajós, donc en bordure d'eau douce. Mais cela n'empêche pas d'essayer de retrouver dans l'échantillon d'image étudié, les zones de végétation proches de l'eau (FIGURE 7.13). Le résultat est intrigant.



(A) (B) (C)

FIGURE 7.13 – Utilisation de la connaissance sur une relation spatiale dans un extrait d'image Landsat à proximité de Santarém (A). En (B), une interprétation automatique sans relations spatiales. Dans l'image (C), les zones de végétation *ExternallyConnected* à des zones d'eau sont représentées en violet.

Certes, les régions détectées répondent à la définition : ce sont bien des zones de végétation *ExternallyConnected* à des zones d'eau. Mais ici, les caractéristiques de l'image sont telles que des représentations de grands espaces boisés sont en contact ou en quasi-contact avec des représentations de surfaces d'eau. Nombre d'entre elles s'enfoncent à l'intérieur des terres plus profondément que d'autres zones de végétation qui elles ne sont pas en contact avec des segments d'eau. De grandes aires végétales intérieures en contact avec des aires aquatiques, alors que d'autres espaces végétalisés plus petits et plus proches du fleuve ne sont pas en contact avec lui : même pour un non-spécialiste, cela peut sembler contradictoire avec ce que l'on pourrait attendre de la définition.

Et en un mot, force est de constater que si l'on cherchait des zones de végétation à proximité de l'eau, alors le résultat ne répond pas vraiment à la question.

Le but ici n'est pas d'apporter une réponse de spécialiste en télédétection, mais de montrer par un exemple qu'une définition doit être pertinente en fonction du concept recherché, du milieu observé et de l'instrument de prise de vue.

L'observation critique d'un résultat doit conduire à s'interroger :

1. La définition décrit-elle bien ce que l'on cherche ?
2. Les caractéristiques de l'image et le vocabulaire disponible permettent-ils de décrire le concept recherché ?
3. Une *bonne* segmentation utilisée pour la recherche de certaines sémantiques est-elle toujours adaptée à la recherche de nouveaux types d'objets dans l'image ?

Autant de questions complexes qui n'ont pas de réponse générique mais qui sont susceptibles de se poser. En l'occurrence on pourrait examiner les pistes suivantes :

1. La définition d'un concept de végétation proche de l'eau n'est peut-être pas pertinente. En particulier, *faute d'expérience suffisante en télédétection*, on ne sait pas si elle correspond à des entités de terrain caractérisées.

2. À supposer que le concept ait du sens, les relations spatiales topologiques ne sont pas forcément les mieux adaptées à sa description.

3. Toujours dans l'hypothèse où le concept aurait un sens, la segmentation utilisée doit être remise en question : si elle ne produit pas d'objets significatifs du point de vue de l'observateur, alors elle ne respecte pas l'hypothèse de *bonne segmentation*.

7.2.4 Potentiel de la parallélisation

Le calcul des relations spatiales objet par objet n'est pas négligeable. Étant donné son coût en $o(n^2)$ (où n est le nombre de segments), il peut être intéressant d'exploiter la parallélisation de l'algorithme. Nous avons ainsi parallélisé ce calcul sur quatre cœurs ce qui réduit considérablement le temps de calcul (TABLE 7.7).

Cœurs \ Objets		1161 objets	5751 objets
1 cœur	(calcul)	$\approx$ 1 h 30 min	$\approx$ 59 h
	(raisonnement)	$\approx$ 3 min	$\approx$ 40 min
4 cœurs	(calcul)	$\approx$ 32 min	$\approx$ 25 h
	(raisonnement)	$\approx$ 3 min	$\approx$ 40 min

TABLE 7.7 – Comparaison de temps d'exécution mesurés avec et sans parallélisation du calcul des relations spatiales. Les 1161 objets sont issus d'un extrait d'image 100×100 pixels. Les 5751 objets sont issus d'un extrait d'image 200×200 pixels. Cette étape fait partie de la description de l'image (instanciation des faits de l'image). Le raisonnement, lui, laissé au raisonneur, n'est pas concerné par la parallélisation.

Il est intéressant de constater que pour les nombres d'entités considérés, l'étape de raisonnement est bien moins coûteuse que le calcul des relation spatiales, en dépit de sa complexité algorithmique bien supérieure.

7.2.5 Influence de la segmentation et des caractéristiques de l'image

Que l'approche soit générale ou étendue par l'intégration des relations spatiales, la relation de l'approche avec la segmentation est duale.

Sur le plan du principe, nous l'avons vu, l'approche est indépendante de la segmentation par hypothèse. La pertinence des résultats exige simplement que la segmentation soit « bonne ». Si cela n'est pas le cas, il existe un risque important que la sémantique de télédétection affectée aux segments soit très éloignée d'une expertise. On pourrait même douter de l'intérêt d'apposer une sémantique sur des groupes de pixel qui n'auraient pas de sens.

Il en va de même des caractéristiques de l'image. L'approche est applicable à toute image satellitaire calibrée en mesures physiques. En revanche, la pertinence des résultats dépend de l'adaptation de la connaissance aux données (contexte de validité). Cela concerne la connaissance dérivée de la radiométrie, de la texture et de la forme des objets. En particulier, la description d'une sémantique de segment est susceptible de varier en fonction des bandes spectrales disponibles.

En ce qui concerne les relations spatiales qui pourraient être interprétées comme proches d'une adjacence, la résolution spatiale peut jouer un rôle déterminant. Sur une image à haute résolution deux représentations d'objets de terrain peuvent apparaître disjointes alors qu'elles se toucheraient sur une image à plus basse résolution, l'espace intermédiaire n'apparaissant pas.

Dans tous les cas, c'est la cohérence de la connaissance avec les données qui est capitale. Les données doivent être compatibles avec le contexte de validité de la connaissance. Ici, ce contexte est défini entre autres par les différentes résolutions, les bandes spectrales disponibles etc.

Résultat 2 (Conclusions de l'intégration des relations topologiques spatiales à l'approche)**.** L'implémentation de l'intégration de relations topologiques spatiales autorise à conclure que :

1. Les relations topologiques spatiales permettent de raffiner les concepts de la connaissance contextuelle.
2. D'un point de vue statistique, l'interprétation ontologique est capable de s'accorder aux dires d'experts de façon tout à fait comparable à l'accord d'experts entre eux.
3. En l'absence de série temporelle, la relation qui semble au premier abord la plus pertinente à considérer entre les objets d'une même image concerne leur potentielle adjacence.
4. L'implémentation du calcul des relations topologiques spatiales peut conduire à des situation contradictoires avec leur sémantique communément admise au premier abord. C'est le cas pour l'interprétation de la relation RCC8 *Externally Connected* comme adjacence.
5. L'intégration des relations topologiques spatiales doit prendre en compte le contexte de validité de la connaissance, en particulier concernant la résolution spatiale.
6. Le temps de calcul des relations topologiques spatiales n'est pas négligeable et la complexité en $o(n^2)$ ne permet à la parallélisation que d'amortir très partiellement l'impact sur les performances.

7.3 Adaptation à la description des pixels

7.3.1 Une déclinaison *a priori* difficile

L'approche objet a mis en évidence les contraintes d'exécution liées à la complexité des logiques de description. Même sur un extrait d'image de taille raisonnable, les temps d'exécution sont importants. Cela pose un problème de passage à l'échelle : la complexité des logiques de description ne laisse rien présager de bon pour une application sur des images entières.

Dans ces conditions il peut sembler étrange de s'être intéressé à la classification des pixels, par définition plus nombreux que les segments. Mais les discussions d'équipes peuvent parfois convaincre d'explorer des voies que l'on aurait autrement négligées.

Il n'en reste pas moins que les premiers résultats ont montré que le raisonnement sur les pixels butait rapidement sur des problèmes de temps d'exécution qui sont apparus bien avant le traitement d'un échantillon d'image significatif (FIGURE 7.14).

En ce qui concerne l'application aux pixels, nous avons utilisé des définitions comparables à celles utilisées pour la classification des objets, pour les pixels de végétation, d'eau et de surfaces minérales ou construites :

$$
\begin{aligned}
VegetatedPixel \doteq & Pixel \\
& \sqcap \exists observedBy.(Observation \\
& \sqcap \exists hasMeasurement.(Measurement \\
& \sqcap \exists hasValue.(PrimitiveValue \\
& \sqcap \exists hasCode.\{double \geq 0,3\} \\
& \sqcap \exists hasCode.\{double \leq 0,8\}) \\
& \sqcap \exists ofCharacteristic.\{ndvi\}))
\end{aligned}
$$

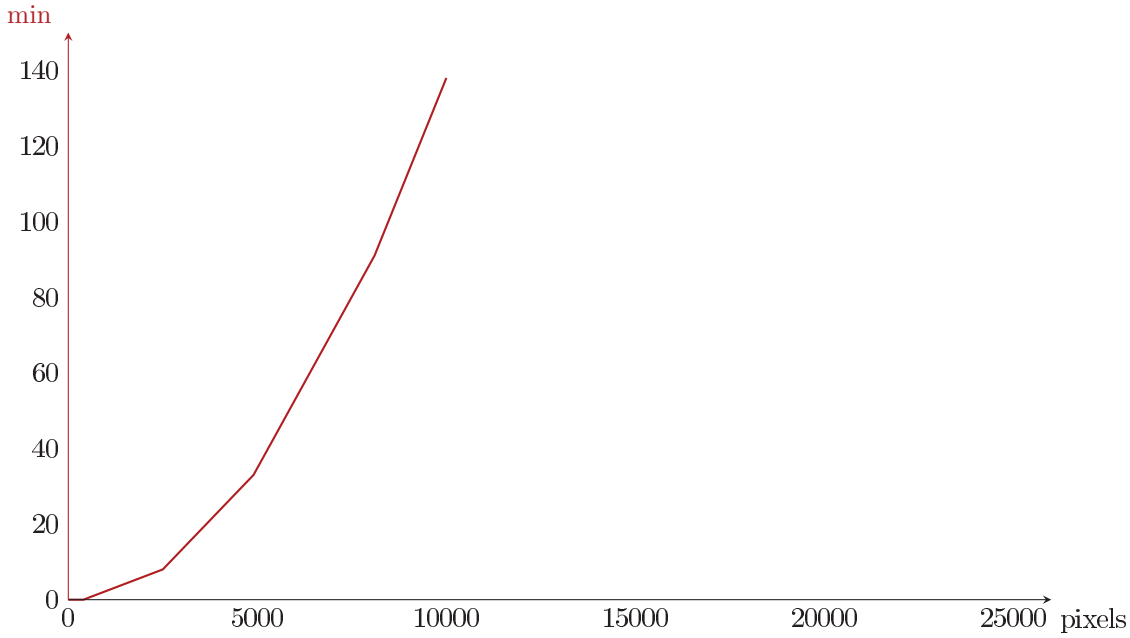


FIGURE 7.14 – Évolution du temps d'exécution du raisonnement sur les pixels en fonction de leur nombre. Cette courbe est basée sur les données de la TABLE 7.8.

$$
\begin{aligned}
WaterPixel \doteq & Pixel \\
& \sqcap \exists observedBy.(Observation \\
& \sqcap \exists hasMeasurement.(Measurement \\
& \sqcap \exists hasValue.(PrimitiveValue \sqcap \exists hasCode.\{double \geq 0,0\}) \\
& \sqcap \exists ofCharacteristic.\{ndwi2\}))
\end{aligned}
$$

$$
\begin{aligned}
MineralPixel \doteq & Pixel \\
& \sqcap \exists observedBy.(Observation \\
& \sqcap \exists hasMeasurement.(Measurement \\
& \sqcap \exists hasValue.(PrimitiveValue \sqcap \exists hasCode.\{double \geq 0,0\}) \\
& \sqcap \exists ofCharacteristic.\{ndbi\}))
\end{aligned}
$$

7.3.2 Contourner la complexité des logiques de description par une hypothèses de travail

Dans une méthode traditionnelle telle que la préclassification proposée par Baraldi et al. (2006), chaque pixel est classifié indépendamment en se basant seulement sur ses caractéristiques intrinsèques. Les relations avec les pixels alentours ne sont pas prises en considération. Rien n'empêche d'appliquer la même hypothèse de travail dans le cadre de leur classification par raisonnement ontologique (Hypothèse 8).

Hypothèse 8 (Hypothèse de la classification des pixels indépendamment les uns de autres)**.** La classification d'un pixel ne dépend que de ses caractéristiques intrinsèques, c'est-à-dire des valeurs spectrales, pseudo-spectrales ou texturales qui lui sont affectées.

Image	Pixels	Temps (A)	(B)
1×1 pixels	1	0	0
2×2 pixels	4	0	0
4×4 pixels	16	0	0
10×10 pixels	100	1 s	0
20×20 pixels	400	11 s	1 s
50×50 pixels	2500	8 min 1 s	7 s
70×70 pixels	4900	33 min 25 s	14 s
90×90 pixels	8100	1 h 31 min 10 s	24 s
100×100 pixels	10000	2 h 18 min 26 s	29 s
150×150 pixels	22500	—	1 min 6 s

Table 7.8 – Comparaison du temps d'exécution du raisonnement sur différents extraits d'image. La colonne (A) correspond à une classification dans une base de connaissances qui contient l'ensemble des pixels d'un extrait. Dans la colonne (B) chaque pixel est classifié dans une base de connaissances distincte.

Cette hypothèse a pour conséquence que la classe d'un pixel ne dépend pas des autres pixels et dans ce cadre, chaque pixel peut donc être classifié séparément.

C'est ce principe que nous avons mis en œuvre dans le but de comparer le comportement de chaque implémentation : classification globale et classification séparée (Table 7.8).

Dans la classification globale, les caractéristiques des pixels dans l'image sont décrites selon le même procédé que les segments d'image lorsque l'approche est appliquée à la suite d'un traitement orienté objet. La description de l'ensemble des pixels de l'image alimente une base de connaissances réalisée ensuite par raisonnement (Table 7.8, colonne (A)).

Dans la classification séparée, le principe de description des pixels est inchangé, mais toutes les descriptions n'alimentent pas une unique base de connaissance. La description du premier pixel alimente une première base, la réalisation est effectuée par raisonnement, le résultat est enregistré, puis la base de connaissances détruite. On répète la même opération pour le second pixel de l'image, et ainsi de suite pour tous les pixels. C'est parce que chaque pixel est classifié séparément des autres que l'hypothèse 8 est nécessaire. Et bien que la technique nécessite de construire de très nombreuses bases de connaissances, le coût de l'opération est linéaire par rapport au nombre de pixels, ce qui se révèle très rapidement rentable en comparaison du coût exponentiel de la classification globale (Table 7.8, colonne (B)).

Les résultats sont sans appel. Alors que la réalisation globale de la base de connaissances subit de plein fouet les conséquences de la complexité des logiques de description, la réalisation individuelle contourne ce problème en linéarisant le temps d'exécution sur la totalité de l'image en fonction du nombre de pixels (Figure 7.15). La pente de la droite résultante correspond au temps approximatif pour classifier un pixel dans la base de connaissance. Elle est donc déterminée principalement par la taille fixe de la TBox et non plus par la taille de la ABox qui dépend du nombre d'entités de l'image décrites.

Il apparaît donc possible d'utiliser la classification par raisonnement au niveau des pixels. On remarquera de plus que rien n'interdit de paralléliser la boucle consistant à classifier les pixels les uns après les autres.

7.3.3 Vers une segmentation ontologique ?

L'application de l'approche aux pixels ouvre des perspectives en ce qui concerne l'approche objet. Deux conséquences se dessinent en effet.

Premièrement, nous avons vu que la méthode de segmentation mise en œuvre dans le cadre de la thèse en amont de la description des segments, se base sur une préclassification des pixels. Il s'agit d'une méthode qui utilise une connaissance incluse dans l'algorithme de Baraldi. Ainsi, dans l'approche objet, les ontologies n'intervenaient que dans la description des objets de l'image, après

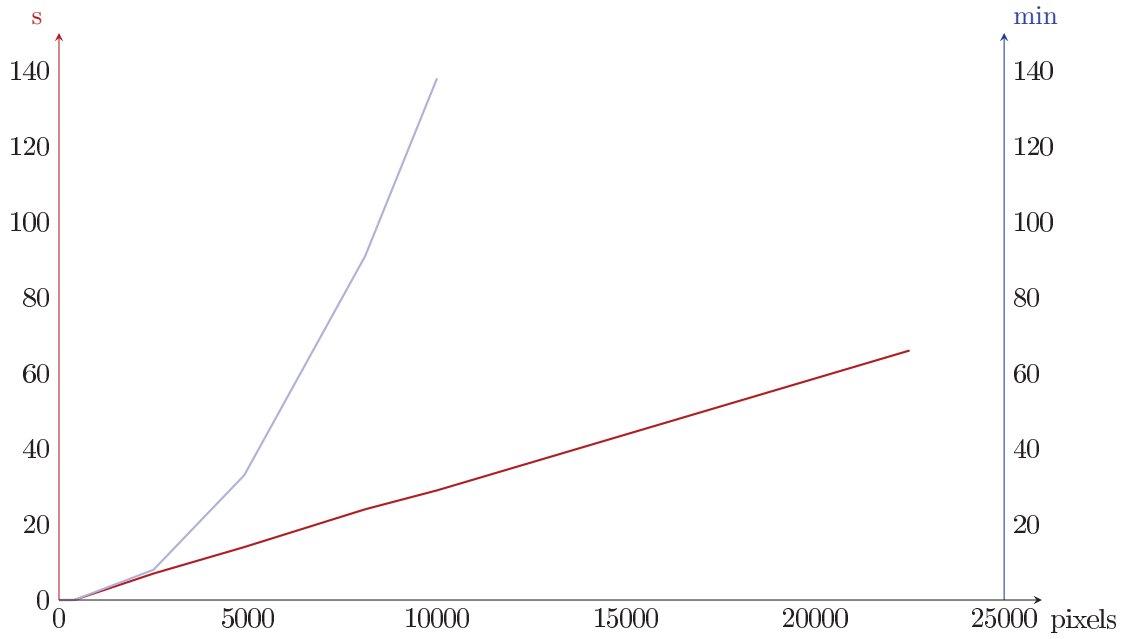


Figure 7.15 – Évolution du temps d'exécution du raisonnement sur les pixels en fonction de leur nombre. La courbe rouge, quasi-linéaire montre le gain de temps d'exécution par rapport au raisonnement sur une base de connaissances globale des faits de l'image (rappelée en bleu). Cette courbe est basée sur les données de la Table 7.8.

la segmentation. L'application aux pixels permet d'envisager d'élargir l'utilisation des ontologies, dans le cadre d'une approche objet, comme préclassification préalable à la segmentation. Si cette combinaison se révélait pertinente, elle permettrait en quelque sorte d'éliminer une préclassification basée sur une connaissance cachée que l'on s'interdit d'utiliser en « l'oubliant » lors de la description, ne gardant de son utilisation que la délimitation des segments.

Deuxièmement, en étendant l'hypothèse 8 à toute entité de l'image (pixels ou segments), on peut appliquer la technique de réalisation séparée à la classification ontologique des objets. Cela exclut cependant la prise en compte des relations spatiales.

Résultat 3 (Conclusions de l'adaptation de l'approche aux pixels)**.** L'implémentation d'une déclinaison pragmatique de l'approche à la classification des pixels permet de conclure que :

1. Contrairement à ce qui pouvait sembler évident, l'utilisation des ontologies pour l'interprétation sémantique des images satellitaires ne concerne pas seulement l'approche par objets, mais également l'approche orientée « pixels ».
2. L'ajustement des hypothèses à la classification indépendante de chaque entité de l'image permet de contourner la complexité des logiques de description par des performances linéaires en fonction du nombre d'entités, pour une connaissance donnée.
3. Les premiers résultats confirment les choix préalables d'appliquer les ontologies en aval de la segmentation. Il ne s'agissait pas d'une hypothèse interdisant définitivement l'utilisation des ontologies en amont.
4. L'examen du rôle des ontologies à différents stades de l'interprétation et du traitement de l'image permet d'imaginer des utilisations de l'approche variées, pragmatiques et complémentaires.

7.4 Classification séparée des objets

7.4.1 Comparaison des classifications globale et séparée

L'effet de la complexité des logiques de description est rapidement perceptible dans l'évolution du temps de raisonnement en fonction du nombre d'entités décrites dans l'image. Déjà contraignant lorsqu'il s'agit d'interpréter des groupes de pixels, le temps d'exécution serait rédhibitoire s'il s'appliquait à une base de connaissances contenant les descriptions de chaque pixel d'une image.

L'astuce de la réalisation de bases de connaissances distinctes pour chaque entité de l'image suppose que leur sémantique puisse être déduite uniquement à partir de caractéristiques intrinsèques. C'est l'hypothèse que nous avons retenue dans le cadre de l'application aux pixels... mais qui pourrait être également intéressante dans une analyse orientée objet.

Cela suppose de n'utiliser dans les descriptions des concepts que des caractéristiques propres à une entité donnée et donc de proscrire l'utilisation de toute *relation*, en particulier spatiale. Comme nous avons pu le voir, l'intégration des relations spatiales pose également des contraintes d'échelle. De plus, bien qu'elle puisse être enrichie par des relations spatiales, l'information radiométrique prend une large part dans la définition des concepts. C'est à partir de cette information intrinsèque que nous avons défini les trois concepts relatifs aux sémantiques que végétation, d'eau et de minéral et de bâti.

Ainsi, le potentiel des relations entre objets pourrait être réservé à une étude locale. Et faire l'hypothèse d'une interprétation de l'image à l'échelle globale, uniquement sur des caractéristiques intrinsèques des objets, semble donc raisonnable.

C'est cette réflexion qui a conduit à adapter à l'analyse orientée objet, l'implémentation utilisée dans le cadre du raisonnement sur les pixels. Les ordres des grandeur des temps d'exécution sont présentés en TABLE 7.9.

Déclinaison de l'approche	Temps du raisonnement
Objet non-optimisée	$>$ 6 h
Pixel optimisée	1 h 36 min
Objet optimisée	10 min

TABLE 7.9 – Comparaison des temps d'exécution de différentes déclinaisons de l'approche.

7.4.2 Produits du raisonnement : du masque sémantique à la carte interprétée

Nous avons vu que selon les choix de combinaisons entre les différents outils, les entrées et les sorties de la chaîne d'interprétation sont susceptibles de changer. L'intégration du raisonnement permet d'obtenir en sortie une collection d'images (des masques) dont chacune correspond à une sémantique de télédétection recherchée.

Voici donc quelques présentations de résultats obtenus à partir de l'image présentée en FIGURE 6.1. Les thématiques sont extraites par raisonnement intégré dans une approche orientée objet en considérant trois sémantiques générales de télédétection : végétation (FIGURE 7.16), minéral/construit (FIGURE 7.17) et eau (FIGURE 7.18).

Ces différents masques sémantiques obtenus par raisonnement peuvent être combinés de manière à obtenir un support assimilable à une carte interprétée de l'image satellitaire (FIGURE 7.19, réalisée semi-automatiquement). Dans le cadre du monde ouvert, trois sémantiques fournissent déjà un résultat intéressant du fait de la possibilité de combiner les concepts. Ces combinaisons peuvent être interprétées de différentes manières.

- De nouvelles sémantiques de télédétection peuvent être définies à partir de sémantiques de télédétection plus élémentaires sans forcément avoir de rapport thématique : Quelle est la signification d'un segment minéral et aquatique ?

- En particulier en périphérie de zones homogènes, les segments peuvent représenter, ou bien des surfaces mixtes, ou bien des surfaces homogènes dont les signaux sont mélangés dans la valeur du pixel. On pourrait ainsi imaginer qu'un segment à la fois minéral et végétal puisse être représentatif, soit d'une surface minérale en partie végétalisée, soit de la frontière entre une surface simplement végétale et une surface strictement minérale.
- La connaissance experte a peut-être été mal exprimée dans l'ontologie de définition des sémantiques de télédétection.
- Enfin, une exception est toujours possible, qui ne remet pas en cause la représentation de la connaissance.

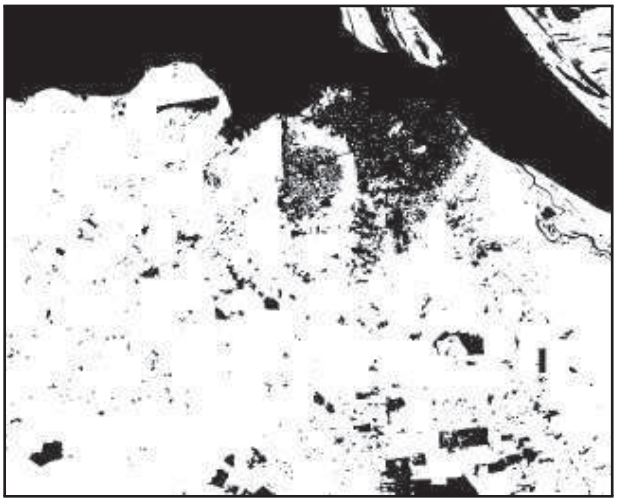


FIGURE 7.16 – Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique végétale. Les segments en noir n'ont pas été classifiés dans cette sémantique.

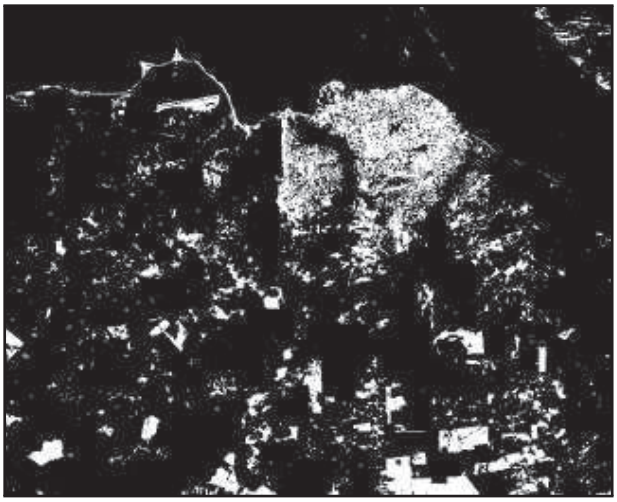


FIGURE 7.17 – Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique minérale ou construite. Les segments en noir n'ont pas été classifiés dans cette sémantique.

FIGURE 7.18 – Extraction, en blanc, des segments de l'extrait (825 × 666 pixels) de l'image Landsat 5 dotés d'une sémantique aquatique. Les segments en noir n'ont pas été classifiés dans cette sémantique.

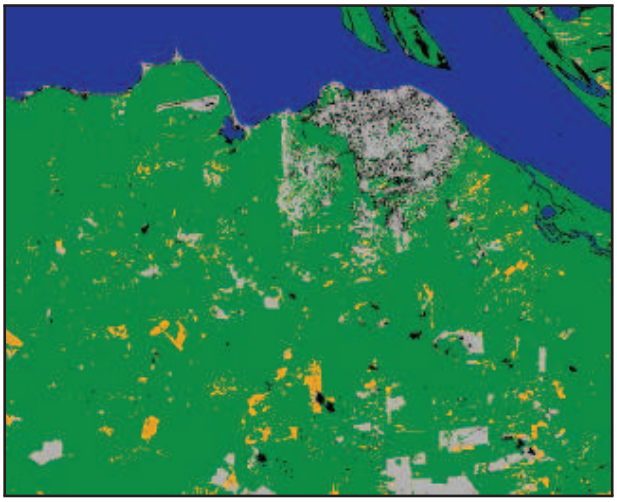


FIGURE 7.19 – Composition colorée des résultats d'extraction des sémantiques végétale (vert), minérale (gris) et aquatique (bleu foncé). D'autres couleurs apparaissent indiquant que certains segments ont été classifiés dans plusieurs classes. Le jaune indique des segments végétaux et minéraux, le bleu clair des segments minéraux et aquatiques. D'un point de vue expert, cela ne signifie pas que ces entités superposent toutes les sémantiques qui leurs sont affectées. Peut-être cela illustre-t-il simplement l'existence d'autres sémantiques comme par exemple des surfaces cultivées, des eaux peu profondes etc.

Résultat 4 (Conclusions de la classification séparée des objets)**.** L'adaptation à la classification des segments, du raisonnement sur des bases de connaissances séparées proposé pour la classification des pixels :

1. Permet d'améliorer significativement la production des résultats.
2. Suppose que la classification d'un objet ne dépende que de ses caractéristiques intrinsèques, ce qui exclut les relations spatiales.

7.5 Réflexions

7.5.1 Valeur ajoutée de l'approche

Une remarque est souvent revenue au cours de la thèse : « *Les résultats ressemblant fortement à ceux produits par une approche classique, quel est l'apport de cette méthode étant donné son handicap en temps de calcul ?* »

Le but de la thèse n'est pas de proposer l'implémentation clef en main d'une approche miracle utilisant les ontologies. La mythification du terme conduit parfois à s'illusionner sur les potentialités qu'il recouvre. L'ambition de la thèse est exploratoire et elle s'appuie sur un changement de paradigme : le passage d'une télédétection assistée par ordinateur à une télédétection guidée par une connaissance experte formalisée, modulaire et partageable.

La ressemblance des résultats présentés avec les produits d'autres méthodes d'exploitation des images est heureuse[1]... mais elle est aussi en partie fortuite. Il est vrai que la détection ontologique des segments de végétation ressemble fortement à un classique seuillage de pixels. Mais, comme nous allons l'expliciter dans l'examen des conséquences de l'hypothèse du monde ouvert, si l'on passe à la négative dans les deux méthodes, le seuillage des pixels indiquera comme un « masque de végétation en négatif », alors que le raisonnement produira une image entièrement noire. Cela illustre que les deux manières de procéder diffèrent profondément : elles reposent en réalité sur les hypothèses opposées du monde ouvert et du monde clos.

Explorer un changement de paradigme alors qu'on en connaît les limitations peut-être risqué, mais cela peut aussi conduire à des résultats inattendus et à des pistes que l'on n'aurait pas soupçonnées.

7.5.2 Les enjeux du passage à l'échelle

L'élargissement de notre approche à des applications complémentaires (intégration des relations spatiales) ou parallèles (application aux pixels) a permis d'illustrer les problèmes de passage à l'échelle liés à la complexité des logiques de description.

La tendance qui se dessine pour le raisonnement global sur les caractéristiques intrinsèques des segments d'image, est aggravée avec la prise en considération des caractéristiques extrinsèques, ne serait-ce que du fait du calcul nécessaire à leur extraction de l'image. C'est ce qui est apparu avec l'intégration des relations spatiales topologiques. Sans en tirer davantage de conclusion, on constatera simplement que la parallélisation ne règle pas le problème des performances du calcul des relations topologiques ; elle se contente de l'amortir à petite échelle.

Mais d'un autre côté, l'application aux pixels a montré que le mur de la complexité pouvait être contourné à défaut d'être abattu. Autant le raisonnement sur l'ensemble des pixels d'une image est inconcevable en l'état actuel de la technologie, autant la classification ontologique d'un pixel après l'autre n'est pas inimaginable, même si elle reste bien moins performante qu'une approche classique.

Il semble donc que les ontologies pourraient être intégrées dans l'exploitation des images bien avant l'étape d'analyse orientée objet. En particulier, l'étape de préclassification préalable à la segmentation pourrait être une bonne candidate à de tels développements.

7.5.3 Vers une approche ontologique hybride pixels/objets ?

Une des deux idées directrices de cette thèse fut de rester centré sur la pierre d'angle du sujet : la jonction entre l'image et la connaissance par le truchement des ontologies. L'autre, fut de rester simple et pragmatique dans l'articulation du traitement d'image avec les ontologies.

Il apparaît à l'issue des pistes qui ont été examinées que le potentiel des ontologies est très riche pour l'exploitation des images satellitaires d'observation de la Terre et qu'elles sont en particulier susceptibles d'intervenir à plusieurs niveaux.

L'approche optimisée sur les pixels peut également être appliquée à la classification des objets, pourvu que leurs caractéristiques soient exclusivement intrinsèques.

[1] Heureusement, les résultats sont cohérents !

Mais dans le cas particulier de l'application que nous avons proposée, l'approche objet repose sur une bonne segmentation elle-même obtenue à partir d'une préclassification des pixels, c'est-à-dire d'une connaissance experte formalisée mais prisonnière du code et de l'algorithme. Dans l'absolu, il est peu probable que cette étape de segmentation produise un jour des résultats probants sans contribution de la connaissance experte. Cela tient de l'intime conviction que la machine ne pourra jamais deviner ce que souhaite obtenir l'expert si celui-ci ne le lui signifie pas. Car de fait, l'attente de chaque expert quant au résultat de la segmentation dépend de *ce qu'il veut voir sur l'image*. Aucune formule mythique, aucun algorithme magique à espérer qui s'adapterait de lui-même d'un expert à l'autre et d'une fois sur l'autre.

Il semble alors que l'on ait tout à gagner à intégrer la connaissance dès avant l'étape de segmentation. Segmenter à partir d'une classification par pixels ne semble pas absurde et nos premiers résultats nous ont conforté dans ce choix.

Le seul problème est l'hétérogénéité de cette connaissance avec la connaissance dans une ontologie. La première est incluse dans le code et l'algorithme, la seconde s'en est libérée. La première est liée à l'hypothèse du monde fermé, la seconde repose sur l'hypothèse du monde ouvert.

À partir du moment où une préclassification des pixels pourrait être obtenue par raisonnement sur une base de connaissances, c'est une grande partie de la problématique amont qui se trouverait résolue par son intégration dans la démarche. Les pixels pourraient alors être préclassifiés selon leurs caractéristiques spectrales et assimilées (on peut inclure les indices et même aller jusqu'à la texture affectée aux pixels), explicitées dans une ontologie.

Le résultat de cette préclassification pourrait ainsi servir de base à l'étape de segmentation. Contrairement à la restriction prise en compte dans notre application pratique, la description de la préclasse des objets produits ne serait pas gênante, bien au contraire. D'une part, son inclusion dans la description est logique puisqu'elle serait elle-même issue d'une ontologie et non plus d'une connaissance cachée. D'autre part, elle pourrait être fort pratique car elle résume en une seule classe, beaucoup de caractéristiques de plus bas niveau (les caractéristiques spectrales). Cela pourrait réduire d'autant les individus peuplant la description de l'image segmentée.

Le raisonnement sur la description des objets serait alors sans doute de ce fait plus efficace et aurait pour objectif d'affiner la préclasse des objets par la prise en compte d'une information à leur échelle : forme, taille, localisation.

Cela vaut pour l'image analysée dans sa globalité, mais la réflexion mérite d'être approfondie aux aspects d'interprétation locale.

Notre avis est que les relations spatiales ont également leur rôle à jouer. La détection des rives a montré leur intérêt. Toutefois, étant donné les limites techniques rencontrées, il peut être pertinent de cantonner leur intervention à l'échelle locale, lorsque peu d'objets sont en jeu. Cela permettrait d'éviter, d'une part un calcul des relations spatiales coûteux en $o(n^2)$, et d'autre part, une explosion du temps de raisonnement qui ne peut plus classifier les objets individuellement dès lors que leur description s'ouvre à des caractéristiques extrinsèques comme les relations spatiales.

7.6 L'hypothèse du monde ouvert

La plupart des méthodes traditionnelles utilisées en télédétection reposent plus ou moins implicitement sur l'hypothèse de monde clos : l'information absente est assimilée à une information fausse. Dans ce schéma de pensée, quand on parcours un arbre de décision, si une entité de l'image est classifiée dans la classe B, alors elle n'est pas membre de la classe A.

Il en va tout autrement des ontologies et des logiques de description qui reposent sur l'hypothèse du monde ouvert, selon laquelle une information absente est potentiellement vraie.

Cette différence joue à deux niveaux : dans la formulation des axiomes terminologiques ou assertionnels d'une part, et dans l'interprétation des résultats d'autre part.

7.6.1 Comment définir les concepts ?

La définition des concepts ne pose pas de problème particulier tant que la négation n'intervient pas.

Prenons l'exemple concret de la définition simplifiée du concept de segment végétal considérée plus haut :

$$\begin{aligned} VegetatedSegment \doteq & Segment \\ & \sqcap \exists observedBy.(Observation \\ & \sqcap \exists hasMeasurement.(Measurement \\ & \sqcap \exists hasValue.(StatisticsSegmentFeature \\ & \sqcap \exists Mean.\{double \geq 0,3\} \\ & \sqcap \exists Mean.\{double \leq 0,8\}) \\ & \sqcap \exists ofCharacteristic.\{ndvi\}\})) \end{aligned}$$

La FIGURE 7.20 rappelle les résultats obtenus avec cette définition utilisée dans le cadre d'une approche objet.

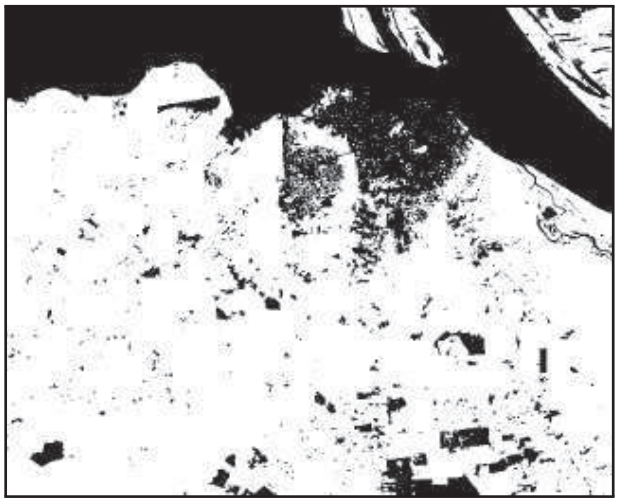


FIGURE 7.20 – Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique végétale. Les segments en noir n'ont pas été classifiés dans cette sémantique.

Comment définir le concept de segment non végétalisé ? Nous avons trois possibilités. La première est de procéder de manière affirmative en décrivant les caractéristiques d'un segment non végétal.

$$NonVegetatedSegment \doteq Segment \sqcap \ldots \tag{7.1}$$

Mais cette manœuvre est hasardeuse, car un segment non végétal peut représenter beaucoup de choses très différentes... Il reste donc la possibilité de définir ce concept par la négation du concept de segment végétal.

$$NonVegetatedSegment \doteq \neg VegetatedSegment \tag{7.2}$$

Enfin, une ultime solution serait d'envisager de définir le concept de segment de végétation comme un segment n'ayant pas les caractéristiques d'un segment végétal. Il s'agit d'une sorte de combinaison entre une définition affirmative et négative. On pourrait par exemple tenter de récupérer des segments qui n'ont pas de mesure de l'indice NDVI comprise entre 0,3 et 0,8 :

$$NonVegetatedSegment \doteq Segment \sqcap \neg(\exists hasMeasurement.Measurement \sqcap \ldots \tag{7.3}$$

Si la première proposition de solution est difficilement définissable, les deux suivantes posent assurément problème à cause de l'hypothèse du monde ouvert.

En posant l'équation 7.2, on pourrait s'attendre en effet à ce que tous les segments n'ayant pas été classifiés comme des segments de végétation soient classifiés dans le concept de non-végétation

FIGURE 7.21 – Résultat auquel on pourrait s'attendre en construisant un concept de *Segment non végétalisé* à partir du concept de *Segment végétalisé* en utilisant le constructeur de négation sur les concepts. En réalité, cette projection de résultat supposerait que tout segment qui n'a pas été classifié comme relatif à de la *végétation* soit relatif à de la *non-végétation*. Cela n'est absolument pas vrai dans le cadre de l'hypothèse du monde ouvert. En revanche, ce type de résultat qui pourrait être obtenu par le parcours d'un arbre de décision.

qui semble complémentaire (FIGURE 7.21). Il n'en est rien, car du fait de l'hypothèse du monde ouvert, un segment qui n'a pas été classifié comme segment de végétation n'est pas pour autant exclu de cette éventualité (FIGURE 7.22).

FIGURE 7.22 – Résultat obtenu si l'on définit un concept de *Segment non végétalisé* à partir du concept de *Segment végétalisé* en utilisant le constructeur de négation sur les concepts. Aucun segment ne répond à l'appel, car du fait de la négation, il ne suffit pas qu'il n'ait pas les caractéristiques d'un *Segment végétalisé*. Il aurait fallu que la description de l'image indiquât de manière explicite ou implicite que les segments en question n'en sont pas dotés. Dans le cas contraire, ils sont toujours susceptibles de l'être, éventuellement dans une autre description.

L'équation 7.3 ne fait que déplacer le problème. On pourrait légitimement espérer qu'un segment qui n'est pas relié à une caractéristique de NDVI requise pour une classification comme segment végétalisé, soit classifié comme segment non végétalisé. Mais là encore, ce n'est pas parce qu'un

segment n'est pas visiblement relié à une telle caractéristique, qu'il n'y est pas relié par ailleurs. Rien ne permet de dire que c'est le cas, mais rien ne permet non plus d'affirmer le contraire.

On le voit, la question de la définition des concepts par l'expert peut poser des problèmes en cas d'utilisation de la négation. Le seul moyen pour contourner le problème est d'affirmer explicitement ou implicitement, qu'un individu *n'est pas une instance de* tel concept ou *n'est pas en relation avec* tel autre individu via une propriété. Falomir et al. (2011) relèvent ce problème et utilise une sorte de description des *non-faits* de l'image... qui peut être très lourde.

7.6.2 Comment décrire un fait ?

La difficulté posée par la négation dans la description de la connaissance experte se pose de la même manière lors de la description des faits, ainsi que le remarque Falomir et al. (2011) dans son exposé des conséquences du monde ouvert. Aucun fait n'est considéré faux s'il n'a été explicitement ou implicitement nié... sans toutefois être pris pour vrai non plus. Or, la plupart du temps nos descriptions sont positives car nous parlons de la présence de ce que nous voyons et non de l'absence de ce que nous ne voyons pas. Ceci est directement lié aux problèmes posés par les définitions du type de l'équation 7.3.

7.6.3 Comment regarder les résultats ?

Cette possibilité d'indétermination des faits oblige à une attitude vis-à-vis des résultats qui diffère du monde clos. Dans un seuillage de pixel ou une classification ordinaire par arbre de décision, une entité qui n'a pas été classifiée dans une classe, n'en est pas membre. Mais dans un résultat obtenu par raisonnement, lorsqu'une entité n'est pas classifiée dans une classe, cela ne veut pas forcément dire qu'elle n'en fait pas partie, mais souvent que l'on n'a pas suffisamment d'information pour l'affirmer.

Ainsi, même dans le cas de résultats apparemment identiques (au sens des données) entre deux approches utilisant le monde clos et le monde ouvert, leur signification (leur sémantique ?) change et ils doivent être regardés différemment.

7.6.3.1 Classes multiples

On pourrait considérer les entités de l'image dotées de plusieurs classes, comme des erreurs, des incohérences, un artefact indésirable. La classification erronée d'une entité reste une possibilité, en particulier si la définition des concepts de télédétection découle d'une connaissance experte inadaptée à l'image. Sauf ce cas qui relève de l'expertise, nous préférons les voir comme une richesse, sans nous prononcer sur leur interprétation liée au domaine, mais simplement en ouvrant quelques pistes de réflexion sur leur signification en faisant justement abstraction du domaine.

En télédétection, les *pixels mixtes* sont toujours source d'interrogations. En particulier, les pixels représentant la limite entre des surfaces de natures différentes reflètent un mélange de leurs caractéristiques spectrales. Les entités de l'image dotés de plusieurs classes peuvent certainement être interprétées de la sorte, soit comme zones frontières, soit comme de larges zones hétérogènes[1]. On constate en effet que la combinaison de classes ne se limite pas à des zones de transition, mais également à des surfaces importantes (FIGURE 7.19, page 147).

Mais on peut également imaginer une autre conséquence à ce résultat : que le résultat de la jonction des caractéristiques des classes soit le reflet d'une nouvelle sémantique de terrain. Cette seconde façon de voir les choses n'est d'ailleurs pas incompatible avec la première, puisque des aires mixtes peuvent aussi être rattachées à des sémantiques qui leurs sont propres, comme certaines aires cultivées etc.

7.6.3.2 Classification supervisée et non supervisée

La définition des classes en intention donne à la classification des entités de l'image une parenté avec les classifications supervisées. Les individus prennent place dans des classes pour lesquelles la définition joue le rôle d'exemples de classification.

[1] On imagine que ce cas de figure pourra être corroboré par certains indices de texture.

Mais en classification supervisée, la jonction des classes prédéfinies n'en crée pas de nouvelles. Au contraire, ici, la combinaison de concepts donne naissance à de nouveaux concepts. Leur interprétation sémantique reste à déterminer... mais il n'empêche que la singularité de leurs caractéristiques est indéniable et parle d'elle-même. En ce sens, la classification ontologique fait penser aux méthodes non supervisées dans lesquelles les classes finales ne sont pas prédéterminées.

Cette méthode de classification est finalement très originale et donne l'impression d'un dialogue entre la connaissance de l'expert d'une part et le contenu de l'image d'autre part. Dans le résultat produit, chacun a son mot à dire : l'expert via les classes qu'il aura définies, l'image par la combinaison surprise qu'elle lui réserve.

Quel que soit l'angle sous lequel on l'aborde, la combinaison des classes illustre que la méthode d'interprétation sémantique proposée est ouverte. Loin de contraindre à une interprétation figée et définitive, elle invite à la réflexion en mettant en lumière les aspects implicites qu'elle aura inférés ou mis en lumière.

Troisième partie

Conclusion

8

Conclusion et ouverture

HORS DES CADRES DE RÉFLEXION QUI ONT JALONNÉ CE DOCUMENT dans la partie de présentation comme dans celle de réalisation, ce chapitre propose quelques considérations générales sur l'ensemble de la thèse. Il s'agit d'abord de tracer un bilan global des axes de recherche explorés, puis d'évoquer quelques enjeux avant de proposer des perspectives de développements ultérieurs.

Sommaire

8.1 Bilan **158**
8.1.1 Résumé structurel 158
8.1.2 Développement de l'approche sur trois axes 158
8.2 Remarques relatives au déroulement des travaux **159**
8.2.1 Risques inhérents au positionnement 159
8.2.2 Communiquer et partager la connaissance 160
8.2.3 Discussion des performances 160
8.2.4 Difficulté de l'interaction avec les experts 161
8.3 Perspectives **161**
8.3.1 En aval : les thématiques de terrain 161
8.3.2 En amont : les prétraitements 163
8.3.3 En profondeur : étiquetage des traitements 164
8.3.4 En largeur : réflexions sur l'identification d'individus composés 164
8.3.5 Dans l'immédiat : enrichir les descriptions des concepts de télédétection 164

8.1 Bilan

8.1.1 Résumé structurel

Nous avons commencé à développer les travaux de thèse dans ce document par la présentation des institutions dans lesquelles a progressivement et naturellement mûri la problématique d'interprétation des images conduite par une connaissance experte.

Une première partie s'est appliquée à poser le contexte disciplinaire des travaux, en portant une attention particulière à désambiguïser le vocabulaire. Il s'agissait tout d'abord de passer en revue les principes de l'imagerie numérique, de la télédétection et des images satellitaires. Puis, ce sont les ontologies et les logiques de description qui ont été examinées.

Un état de l'art de l'utilisation de la connaissance formalisée dans l'exploitation des images a ensuite permis de se rendre compte de la variété des propositions (segmentation, modélisation de la connaissance, recherche sémantique...) et des domaines d'application (photographies, télédétection, médecine...).

L'approche générale présentée en deuxième partie, se base sur le principe d'interprétation par raisonnement d'abord exploré par Schober et al. (2004). Ce principe a été repris par Falomir et al. (2011) et actualisé avec le développement des langages du web sémantique. C'est ce principe d'interprétation qui a guidé notre approche en télédétection.

Cette approche a été mise œuvre avec trois préoccupations : sémantique, temporelle et spatiale.

8.1.2 Développement de l'approche sur trois axes

La préoccupation sémantique constitue l'axe principal du travail de thèse. Elle a consisté à utiliser les ontologies dans un but d'interprétation des images satellitaires automatisable, en réduisant le fossé sémantique, quel que soit le paradigme d'analyse (pixel ou objet).

Les autres préoccupation correspondent à la prise en compte de deux axes secondaires. Un axe *méthodologique* concernant l'espace, par l'examen des relations topologiques spatiales. Et un axe *applicatif* concernant le temps d'exécution, par l'examen d'hypothèses et de techniques qui ont permis de contourner la complexité des logiques de description.

8.1.2.1 Le fossé sémantique

Nous l'avons vu, la limite entre la connaissance et le traitement n'est pas toujours facile à établir. De ce point de vue, l'approche proposée se veut pragmatique. Elle est d'autant plus proche de l'image, qu'elle en donne une description quantitative, par opposition aux propositions de Schober et al. (2004) et Falomir et al. (2011) qui passent par un modèle qualitatif. Mais notre approche ne donne pas pour autant aux ontologies l'exclusivité de la connaissance. Les calculs arithmétiques tels les indices spectraux sont laissés au traitement d'image, même lorsqu'ils amorcent un lien avec une connaissance.

La réduction du fossé sémantique repose sur l'utilisation des ontologies dans plusieurs descriptions explicites.

1. Tout d'abord, la description des conceptualisations élémentaires de chaque acteur (utilisateur humain ou brique logicielle), et des correspondances qui les relient. Les conceptualisations de référence de tous les acteurs, reliées entre elles, forment la conceptualisation de référence du système.

2. Ensuite, la conceptualisation de référence du côté de la brique logicielle sert à la description de l'image et des entités qui la composent (pixels ou segments). Elle rend compte d'une connaissance sur le contenu de l'image : les faits de l'image.

3. Enfin, la conceptualisation de référence du côté utilisateur sert à la description conceptuelle de la connaissance experte. Elle rend compte d'une expérience sur la manière dont les entités sont interprétées dans une sémantique de télédétection, en fonction de leurs caractéristiques dans l'image.

L'ensemble de ces descriptions, appuyé sur les logiques descriptives, peut être exploitée par des logiciels capables d'inférer une connaissance implicite. C'est avec de tels raisonneurs qu'apparaît explicitement l'interprétation sémantique des entités de l'image, grâce à l'expérience formalisée par l'expert.

8.1.2.2 Le temps

Les séries temporelles n'ont pas été étudiées dans le cadre de cette thèse. Chaque description concerne une unique image et n'utilise pas de relation temporelle entre les entités.

Par la considération du temps, il s'agissait d'examiner un point capital de l'approche : la complexité algorithmique du raisonnement, induite par les logiques de description, est croissante en fonction de leur expressivité.

La complexité des logiques de description a été étudiée dans de nombreux cas et laisse peu d'espoir d'obtenir des résultats dans des temps raisonnables sur des bases de connaissances représentant des images volumineuses.

Il est apparu intéressant de proposer une manière de contourner la question, plutôt que d'affronter directement un problème de complexité apparemment irréductible. Dans le cas où les entités de l'image peuvent être interprétées uniquement sur leurs caractéristiques intrinsèques, nous avons proposé de les classifier dans des bases de connaissances séparées. Cette astuce permet de linéariser le temps d'exécution en fonction du nombre d'entités de l'image.

Le temps d'exécution pouvait initialement paraître très restrictif voire dissuasif. Il s'avère en fait possible de l'améliorer très substantiellement, à condition de renoncer à toute classification des entités de l'image sur des caractéristiques extrinsèques.

8.1.2.3 L'espace

Pourtant, un type de caractéristiques extrinsèques a été exploré : les relations topologiques spatiales entre les segments d'image.

Comme concepts ou rôles élémentaires potentiellement utilisés dans les descriptions de l'image numérique et de la connaissance experte, les relations spatiales doivent être définies dans la conceptualisation de référence. Nous avons fait le choix d'une intégration simplifiée, sur le modèle spécifié par l'ontologie GeoSPARQL.

Parce qu'elles interdisent le recours à l'astuce d'amélioration du temps d'exécution, les relations spatiales sont rédhibitoires à l'échelle globale. On constate pourtant que malgré leurs limites, elles sont susceptibles d'enrichir l'interprétation sémantique de l'image grâce à la définition de nouveaux concepts. La sémantique qu'il a semblé le plus pertinent d'étudier renvoyant à l'idée d'adjacence, il se peut que l'intérêt des relations spatiales réside en leur utilisation à l'échelle locale. On peut espérer à cette échelle réduire le nombre d'objets considérés dans une même base de connaissances, tout en ayant accès aux relations spatiales témoignant de leur proximité.

8.2 Remarques relatives au déroulement des travaux

8.2.1 Risques inhérents au positionnement

La thèse se plaçant à la charnière entre ingénierie des connaissances et télédétection, ces deux domaines nécessitaient une présentation relativement approfondie. La principale caractéristique de l'association entre les ontologies et les images satellitaires concerne la multiplicité des approches possibles en fonction des objectifs (fouille de données, segmentation...). Un tour d'horizon des propositions formulées à ce jour a été proposé dans l'état de l'art.

En ce qui nous concerne, la question tournait principalement autour de l'interprétation sémantique des images : comment réduire le fossé sémantique ?

Nous avons choisi une approche exploitant les mécanismes de raisonnement basés sur les logiques de description. Malgré les limitations inhérentes aux raisonneurs (particulièrement les questions liées à la complexité des logiques descriptives), ce choix semblait intéressant car il permettait d'exploiter le plus possible les conséquences de la formalisation de la connaissance. Et puisque la

formalisation présente l'inconvénient d'appauvrir l'information, autant essayer de profiter le plus possible de ses avantages !

Ce choix présentait le risque de mener à une impasse, même si les travaux de Falomir et al. (2011) sur d'autres images laissaient entrevoir des possibilités intéressantes. Mais quand bien même toute piste se serait révélée sans issue, il aurait été intéressant d'explorer cette voie.

En fait, l'approche mène bien à une impasse. Les performances sont là qui témoignent des limites du principe. Mais cette impasse n'est pas totalement fermée, puisque malgré le peu de concepts experts, des résultats déjà intéressants ont été obtenus. C'est d'ailleurs l'application aux pixels qui a apporté le plus de surprises, alors qu'elle aurait facilement pu être négligée. Mais nous résumerons un peu plus loin, les conclusions relatives aux performances...

8.2.2 Communiquer et partager la connaissance

Le principal avantage de l'approche réside en sa modularité. À l'expert la connaissance, à la machine le traitement : les ontologies servent d'intermédiaire. Grâce aux ontologies et à la standardisation, la connaissance peut être séparée de l'algorithme, puis du code, à condition de le faire avec juste mesure, en laissant les calculs arithmétiques à l'extérieur des ontologies.

Mais les avantages peuvent aussi être plus subtils et se mêler à des inconvénients. L'approche autorise la classification des entités dans plusieurs classes, ce qui évite de biaiser les conséquences de l'expression de la connaissance par l'expert. C'est aussi une vision des choses créative qui pousse à se poser des questions sur le produit de la connaissance experte et la sémantique de classes combinées.

L'hypothèse du monde ouvert, qui n'est pas toujours évidente à manier, a une influence sur la manière d'exprimer la connaissance et de considérer les résultats. Elle se contente souvent de dire : « je ne sais pas si cela en est », là où le monde clos se serait définitivement prononcé : « cela n'en est pas ». Cela n'est pas étonnant d'une approche dont la finalité est de favoriser le partage de la connaissance, qu'elle repose sur une grande ouverture qui s'impose jusque dans la manière dont on doit appréhender ses résultats.

Enfin, l'ensemble du travail de mise en œuvre a été réalisé avec des outils logiciels sous licence libre : OTB, FaCT++, Protege, owl_cpp et éventuellement Soprano. Ce choix a permis de collaborer avec le développeur de owl_cpp, M. LEVIN.

D'autres collaborations ont été mises en place à l'intérieur de l'unité de recherche, afin d'examiner le potentiel des ontologies en télédétection, en particulier pour la représentation de nomenclatures des entités de terrain.

8.2.3 Discussion des performances

Commençons par les performances en temps d'exécution. En ce qui le concerne, quelle que soit sa déclinaison, l'implémentation de notre approche ne peut concurrencer les algorithmes des logiciels de télédétection assistée par ordinateur.

On sent bien intuitivement qu'un logiciel de raisonnement logique, indépendant de la connaissance et basé sur le monde ouvert n'est pas comparable à un arbre de décision, construit pour un problème spécifique et s'appuyant sur le monde clos. Dans le premier cas, on prend en considération le plus de portes possibles. Dans le second cas on les ferme dès que l'occasion se présente. « Dans les *systèmes à base de connaissances*, [...] bien que nettement séparés, la représentation de la connaissance et les mécanismes d'exploitation de cette connaissance sont interdépendants : un raisonnement adapté à une représentation peut s'avérer lourd et compliqué pour une autre » (Mariño Drews, 1993).

Il est important toutefois de remarquer que l'utilisation des ontologies dans les images satellitaires à des fins de raisonnement, aurait pu sembler très improbable eu égard à la quantité de données d'une part, et à la complexité des logiques de description d'autre part. Pourtant, certaines application de l'approche ont démontré qu'elle pouvait produire des résultats dans des délais très acceptables, à condition de restreindre le nombre des possibles par hypothèse (classification sur les caractéristiques intrinsèques) et d'en tirer les conséquences dans l'implémentation (classification dans des bases de connaissances séparées). La prise en compte de connaissances sur les relations

spatiales, qui ne permet pas d'utiliser de telles astuces, reste problématique de par son coût en temps d'exécution élevé, même à l'échelle locale.

Mais raisonner vite ne sert à rien si l'on raisonne faux. Nous avons vu que les raisonneurs souvent validés sur de petites bases de connaissances, n'offrent pas vraiment de garantie quant à la validité de leurs résultats sur des bases volumineuses et complexes. Il n'est pas évident de « valider » les résultats concrets de l'approche.

Pour les sémantiques retenues dans le cadre des essais applicatifs, nous avons confronté le raisonnement sur des objets de l'image à un seuillage de pixels. Les résultats ont été comparés qualitativement (images similaires) et quantitativement (matrices de confusion et indices statistiques). En ce qui concerne les relations spatiales, ce sont des dires d'experts qui ont été utilisés et qui ont montré que le raisonnement ontologique a toute sa place en télédétection, même s'il ne doit pas être mythifié.

8.2.4 Difficulté de l'interaction avec les experts

Quoique le principe ait pu être expérimenté à partir de quelques concepts de télédétection, l'expression de la connaissance experte reste un facteur limitant très important. Même si elle ne représentait qu'un aspect annexe de la thèse, l'interaction avec les experts s'est avérée complexe. L'élicitation de la connaissance est une tâche ardue qui ne peut être envisagée que si l'informaticien et le télédétecteur convergent ensemble vers une compréhension mutuelle. Or, cette action commune ne va pas de soi, ne serait-ce que parce que l'informaticien et l'expert en télédétection ont chacun, ses propres références, son propre vocabulaire, ses propres analogies qui lui parlent.

Ni l'un ni l'autre n'a en général la capacité de faire la totalité du chemin, à moins de cumuler les deux compétences. Mais l'idéal est tout de même qu'ils s'investissent chacun dans le domaine de l'autre, de manière à apprendre à se comprendre. À partir de là, des discussions peuvent s'engager, qui au-delà d'exprimer le contenu d'une expérience, se révèlent porteuses d'idées enrichissantes. C'est ainsi en tout cas que nous avons pu travailler.

8.3 Perspectives

L'élaboration de la proposition et la réflexion sur la mise en œuvre ont été concrétisées par plusieurs déclinaisons implémentées et appliquées à une zone d'étude.

Cette thèse s'inscrit dans le cadre du projet CARTAM-SAT qui concerne l'ensemble du processus de télédétection, depuis l'acquisition des images jusqu'à la diffusion de la connaissance sous toutes ses formes (cartes interprétées et enseignements). Il est donc logique d'envisager quelques perspectives de manière à donner une ouverture à nos réflexions et propositions méthodologiques et techniques.

Ces perspectives s'inscrivent dans une vision globale d'une chaîne de traitements paramétrée par ontologies.

8.3.1 En aval : les thématiques de terrain

Le projet CARTAM-SAT, comme le projet BIO_SOS ont une grande envergure et englobent avec l'image satellitaires, d'importantes perspectives d'application à diverses thématiques de terrain (Arvor et al., 2012; Andrés et al., 2012a).

Une réflexion sur les nomenclatures LCCS et GHC a été explorée par Arvor et al. (2013a). Une partie importante de ces nomenclatures a été mise sous forme d'ontologies, avec la possibilité de passer de l'une à l'autre. Ce travail et la thèse se complètent harmonieusement en permettant aux ontologies de couvrir l'ensemble des domaines de l'image satellitaire (structures et traitements) jusqu'aux entités de terrain (LCCS et GHC) en passant par l'interprétation des entités de l'image en télédétection. Peut-être qu'une mise en correspondance entre la connaissance experte utilisée pour l'interprétation de l'image d'une part, et les nomenclatures de terrain d'autre part, permettrait-elle de transformer une image satellitaire en carte interprétée faisant référence à une nomenclature de terrain ? (FIGURE 8.1)

OBOE n'est pas à proprement parler une ontologie de modélisation d'entités, mais de modélisation de l'*observation* d'entités. Ce choix, qui peut être discuté, complexifie la modélisation. Mais

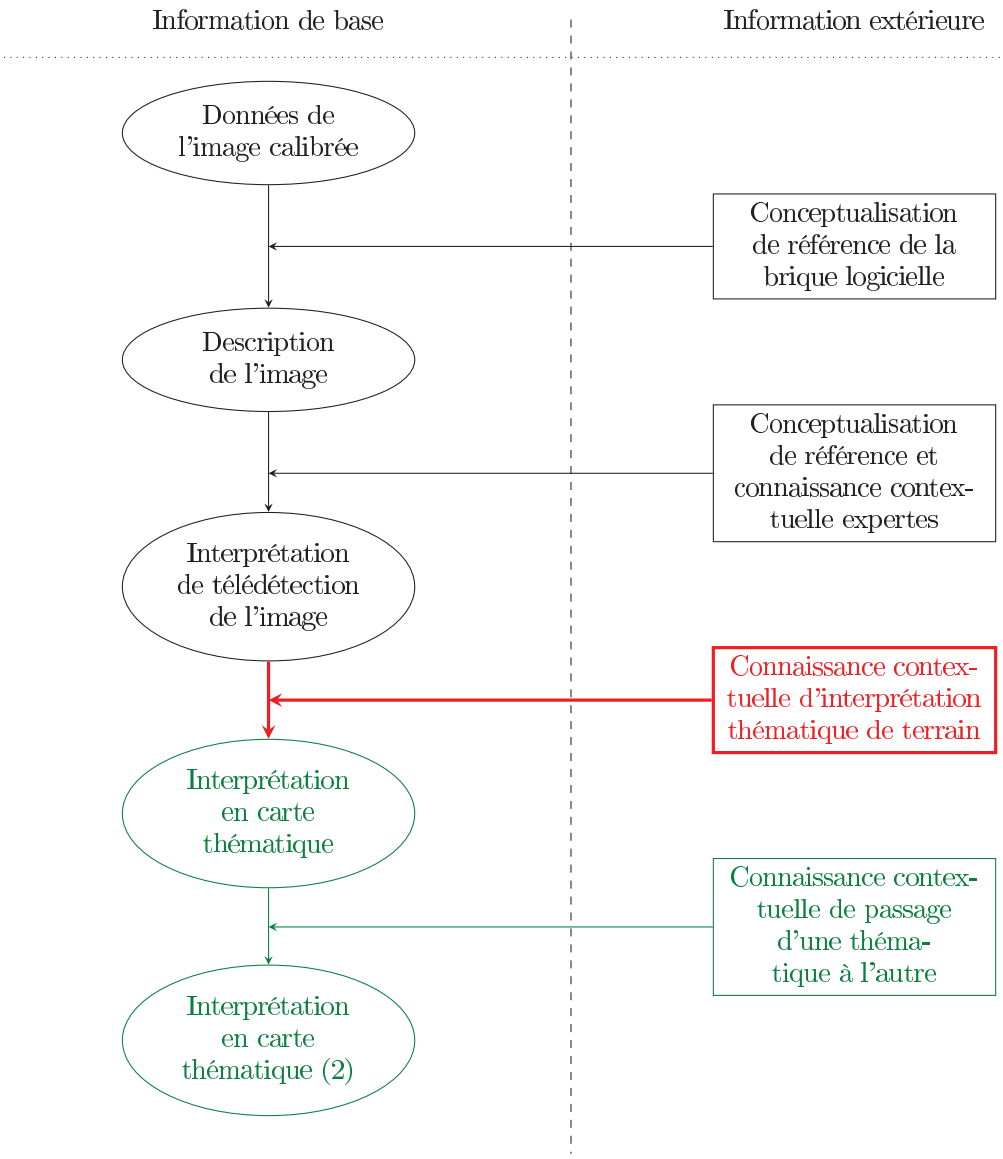


FIGURE 8.1 – Extension envisagée de l'approche en aval de l'interprétation de l'image, vers des cartes thématiques. La partie explorée au cours de la thèse (en noir) et celle explorée par Arvor et al. (2013a) (en vert) sont susceptibles d'être reliées (en rouge) par la formalisation dans une ontologie de la connaissance nécessaire pour passer d'une interprétation des entités de l'image en classes d'occupation du sol à des représentations thématiques.

il encourage aussi une vision des choses par couches d'observation. De la même manière que nous avons considéré les caractéristiques des entités de l'image comme étant le fruit d'observations et de mesures, peut-être pourrions-nous considérer le segment lui-même comme la valeur d'une mesure relative à l'observation d'une entité de terrain et dont la caractéristique ne serait autre que le capteur utilisé pour réaliser la mesure ? (FIGURE 8.2)

Enfin, tant qu'à manipuler les ontologies, peut-être serait-il possible de les utiliser à des fins de présentation, par exemple pour paramétrer la couleur d'une classe sur une carte interprétée.

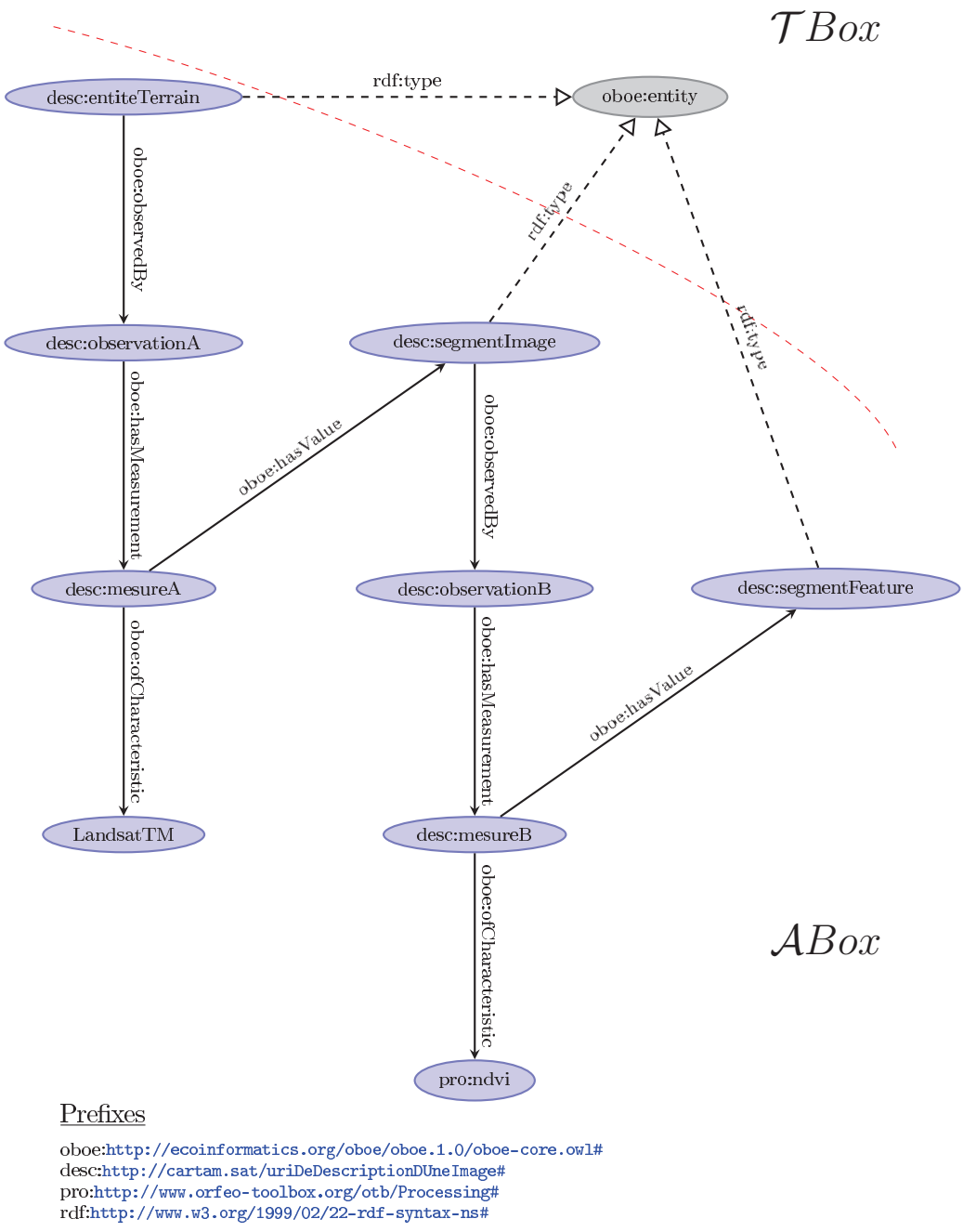


FIGURE 8.2 – Une possibilité de liaison entre les descriptions des objets de l'image et celles des objets de terrain pourrait consister à considérer que les objets de l'image sont le fruit d'observations d'objets de terrain en utilisant l'ontologie OBOE. Alors que l'observation d'un objet de l'image peut être caractérisée par un traitement, l'observation d'un objet de terrain pourrait être caractérisée par le capteur utilisé...

8.3.2 En amont : les prétraitements

Les métadonnées qui accompagnent les images satellitaires sont prises en compte par les prétraitements, pour la production de l'image calibrée qui nous a servi de donnée initiale.

Nous avons vu que notre proposition, initialement destinée à s'appliquer à une approche objet, a finalement trouvé une application intéressante en classification des pixels. Si cette classification pouvait être utilisée en lieu et place de la classification de Baraldi et al. (2006) comme base de

segmentation, alors cela constituerait une extension de l'approche en amont de l'interprétation des objets.

Une question se pose alors : ne pourrait-il pas être pertinent de prolonger encore cette piste en amont et d'explorer le potentiel de la représentation des métadonnées de l'image dans une ontologie ? Pourrait-on alors proposer un mécanisme de calibration des images paramétré par ontologie ? C'est peu probable, étant donné qu'il s'agit d'opérations arithmétiques[1]. Mais cela vaut peut-être la peine de se poser la question : comme dans le cas de la classification des pixels, on peut avoir des surprises moyennant quelques adaptations.

La FIGURE 8.3 illustre certains rôles que les ontologies pourraient jouer dans une chaîne de traitements idéalement paramétrée par de la connaissance formalisée.

8.3.3 En profondeur : étiquetage des traitements

Les ontologies apportent une certaine modularité, mais les applications expérimentales implémentées dans le cadre de la thèse manquent encore de souplesse. En effet, l'étape de description de l'image oblige à opérer au préalable l'ensemble des traitements utilisés comme caractéristiques dans la description de la connaissance experte.

Cela produit un code source quelque peu rigide qui contraste avec l'idée de modularité. Supposons que la connaissance experte évolue : si elle utilisait moins de traitements, alors certains seraient désormais calculés inutilement. Mais si, au contraire, elle en utilisait davantage : il faudrait alors modifier le code. Pour assouplir tout cela, il pourrait être intéressant de n'opérer sur les données d'entrée *que* les traitements utilisés par la connaissance experte. Il semblerait que dans cette tâche, le raisonnement ontologique puisse également être utile.

1. Dans un premier temps, il faudrait extraire de la connaissance experte, tous les traitements utilisés dans la définition d'un concepts doté d'une sémantique de télédétection.
2. Dans un deuxième temps, adapter dynamiquement la chaîne de traitements de manière à ce que ne s'exécutent que les traitements nécessaires.

Cette perspective de travail pourrait peut-être tirer profit des propositions de Lin (2011) pour la constitution de chaînes de traitements.

8.3.4 En largeur : réflexions sur l'identification d'individus composés

Nous avons évoqué dans les discussions sur l'éventuelle représentation de relations temporelles, que les ontologies ont un caractère profondément statique. De même, les ontologies ne *créent* pas d'entités dynamiquement : elles *décrivent* seulement ce qui est.

C'est une question qui aurait pu se poser dans une certaine mesure, dans la perspective d'utilisation des ontologies pour guider une segmentation. Nous avons évoqué cette piste sous l'angle d'une classification préalable à la segmentation proprement dite. Car la création des segments comme entités nouvelles, ne peut être le fruit d'un raisonnement. Elle dépend d'un regroupement des pixels en de *nouvelles instances* de segments d'image. Ce regroupement est le fruit d'un traitement spécifique.

De la même manière qu'un segment peut être défini comme composé de pixels, on peut se poser la question de la constitution de groupes de segments pour représenter de nouvelles entités complexes. C'est un problème qui intervient par exemple dans la description des paysages. Et sans apporter de réponse catégorique à ce type de situation, il nous semble essentiel d'utiliser les ontologies de manière pragmatique, en association avec d'autres techniques, comme nous l'avons proposé avec le traitement d'image.

8.3.5 Dans l'immédiat : enrichir les descriptions des concepts de télédétection

On le voit, les perspectives sont variées qui pourraient faire l'objet de travaux spécifiques. Mais il y a un point qui mériterait d'être exploré non pas tant pour étendre, mais pour étoffer l'approche que nous avons proposée : enrichir les descriptions des sémantiques de télédétection.

[1] Nous avons vu que les ontologies sont mieux adaptées aux opérations logiques qu'arithmétiques.

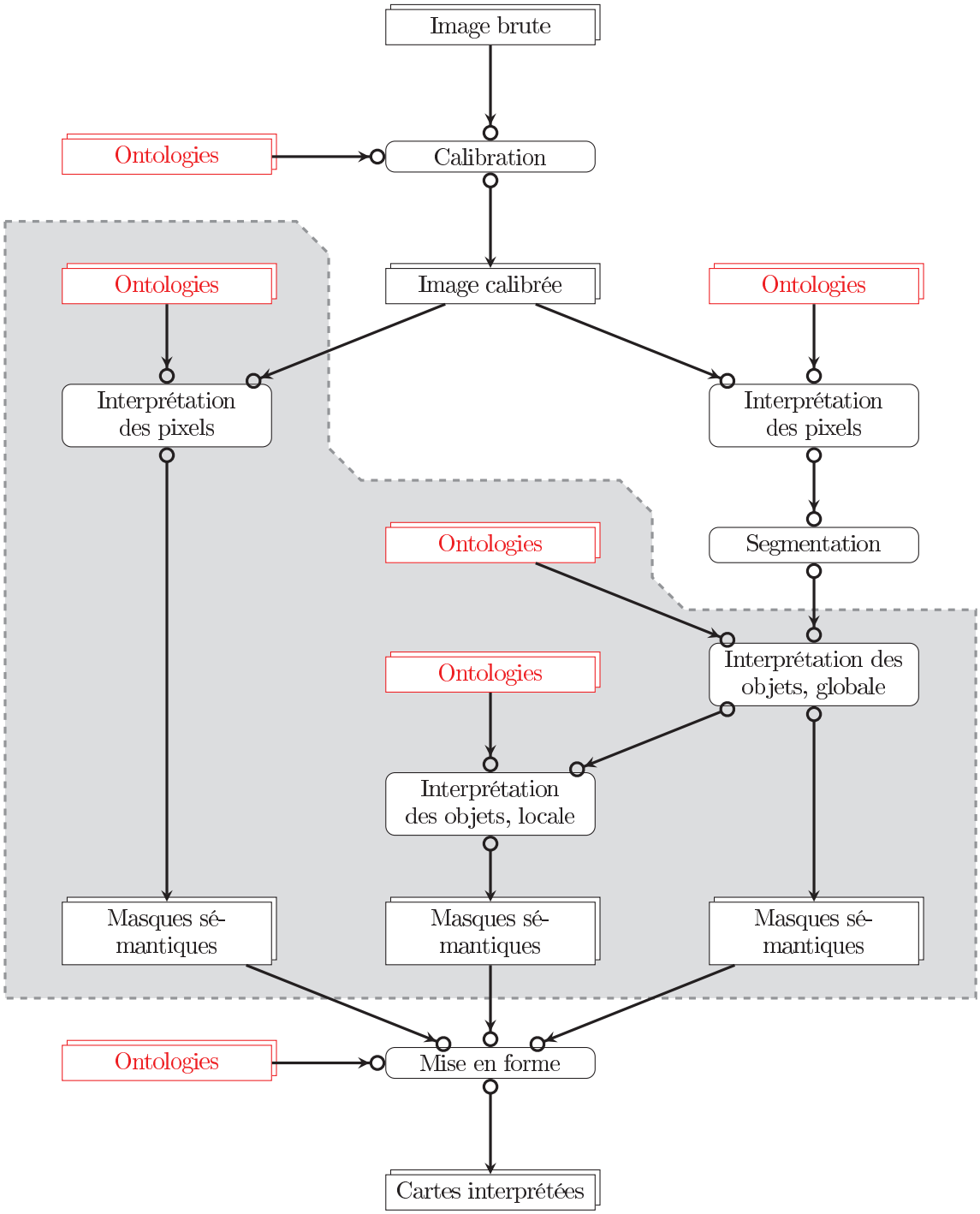


FIGURE 8.3 – Récapitulatif de quelques perspectives en amont, représenté sous la forme d'une chaîne de traitements. La thèse s'est penchée sur l'exploration des domaines grisés.

Nous nous sommes appuyés sur trois concepts de base pour illustrer notre proposition et obtenir des résultats concrets. Mais de nombreux autres concepts pourraient être définis, ne serait-ce que sur des caractéristiques intrinsèques, pour obtenir des interprétations sémantiques plus détaillées.

Il est nécessaire pour cela d'explorer davantage les sources évoquées de connaissance experte. Les travaux de Arvor et al. (2013b) et Bayoudh et al. (2012) qui ont été mentionnés pourraient sans doute concourir à cet objectif.

Quatrième partie

Annexes

A

Quelques indices utilisés en traitement d'image

A.1 Indices spectraux

La TABLE A.1 recense quelques indices spectraux implémentés dans *Orfeo Toolbox*[1]. L'interface de programmation permet également de composer des indices personnalisés.

Nom	Signification	Documentation de référence
ARVI	*Atmospherically Resistant Vegetation Index*	Kaufman and Tanre (1992)
EVI	*Enhanced Vegetation Index*	Huete et al. (1994)
IB	*Brilliance Index*	Détection de sols avec les canaux vert et rouge.
IB2	*Brilliance Index 2*	Détection de sols avec les canaux vert, rouge et proche infrarouge.
IC	*Color Index*	Pouget et al. (1990)
IPVI	*Infrared Percentage Vegetation Index*	Crippen (1990)
IR	*Redness Index*	Pouget et al. (1990)
ISU	*Index surfaces built*	Abdellaoui and Rougab (1997)
MNDWI	*Modified Normalized Difference Water Index*	Xu (2006)
MSAVI	*Modified Soil Adjusted Vegetation Index*	Qi et al. (1994)
MSAVI2	*Modified Soil Adjusted Vegetation Index 2*	Qi et al. (1994)
NDBI	*Normalized Difference Built Up Index*	Zha et al. (2003)
NDPI	*Normalized Difference Pond Index*	Lacaux et al. (2007)
NDTI	*Normalized Difference Turbidity Index*	Lacaux et al. (2007)
NDVI	*Normalized Difference Vegetation Index*	Rouse et al. (1973)
NDWI	*Normalized Difference Water Index*	Gao (1996)
NDWI2	*Normalized Difference Water Index 2*	McFeeters (1996)
PVI	*Perpendicular Vegetation Index*	Richardson and Weigand (1977)
RVI	*Ratio Vegetation Index*	Pearson and Miller (1972)
SAVI	*Soil Adjusted Vegetation Index*	Huete (1988)
SRWI	*Simple Ratio Water Index*	—
TNDVI	*Transformed NDVI*	—
TSARVI	*Transformed Soil ARVI*	Kaufman and Tanre (1992)
TSAVI	*Transformed Soil Adjusted Vegetation Index*	Baret and Guyot (1991)
WDVI	*Weighted Difference Vegetation Index*	Clevers (1988)

TABLE A.1 – Quelques indices spectraux implémentés dans *Orfeo Toolbox*.

[1] http://www.orfeo-toolbox.org/doxygen/namespaceotb_1_1Functor.html

A.2 Indices de texture

Les quelques formules explicitées ici permettent de calculer la valeur de différentes textures d'un pixel p à partir de sa matrice de co-occurrences en niveaux de gris ($GLCM$) dont les $g(i,j)$ représentent les composantes de la i^{e} ligne et j^{e} colonne. Ces indices sont implémentés dans *Orfeo Toolbox*[1]. D'autres ont été proposés par Haralick et al. (1973).

Energy

$$Energy(p) = \sum_{i,j}^{GLCM(p)} g(i,j)^2$$

Entropy

$$Entropy(p) = -\sum_{i,j}^{GLCM(p)} g(i,j)\log_2(g(i,j))$$

Correlation

$$Correlation(p) = -\sum_{i,j}^{GLCM(p)} \frac{(i-\mu)(j-\mu)g(i,j)}{\sigma^2}$$

Avec $\mu = \sum_{i,j}^{GLCM(p)} i.g(i,j)$ et $\sigma = \sum_{i,j}^{GLCM(p)}(i-\mu^2)g(i,j)$.

Difference Moment

$$DifferenceMoment(p) = -\sum_{i,j}^{GLCM(p)} \frac{1}{1+(i-j)^2}g(i,j)$$

Inertia (Contrast)

$$Inertia(p) = -\sum_{i,j}^{GLCM(p)} (i-j)^2 g(i,j)$$

Cluster Shade

$$ClusterShade(p) = \sum_{i,j}^{GLCM(p)} ((i-\mu)(j-\mu))^3 g(i,j)$$

Cluster Prominence

$$ClusterProminence(p) = \sum_{i,j}^{GLCM(p)} ((i-\mu)(j-\mu))^4 g(i,j)$$

[1] http://www.orfeo-toolbox.org/doxygen/classotb_1_1HaralickTexturesImageFunction.html

B

Essais d'utilisation des relations spatiales réifiées d'OBOE

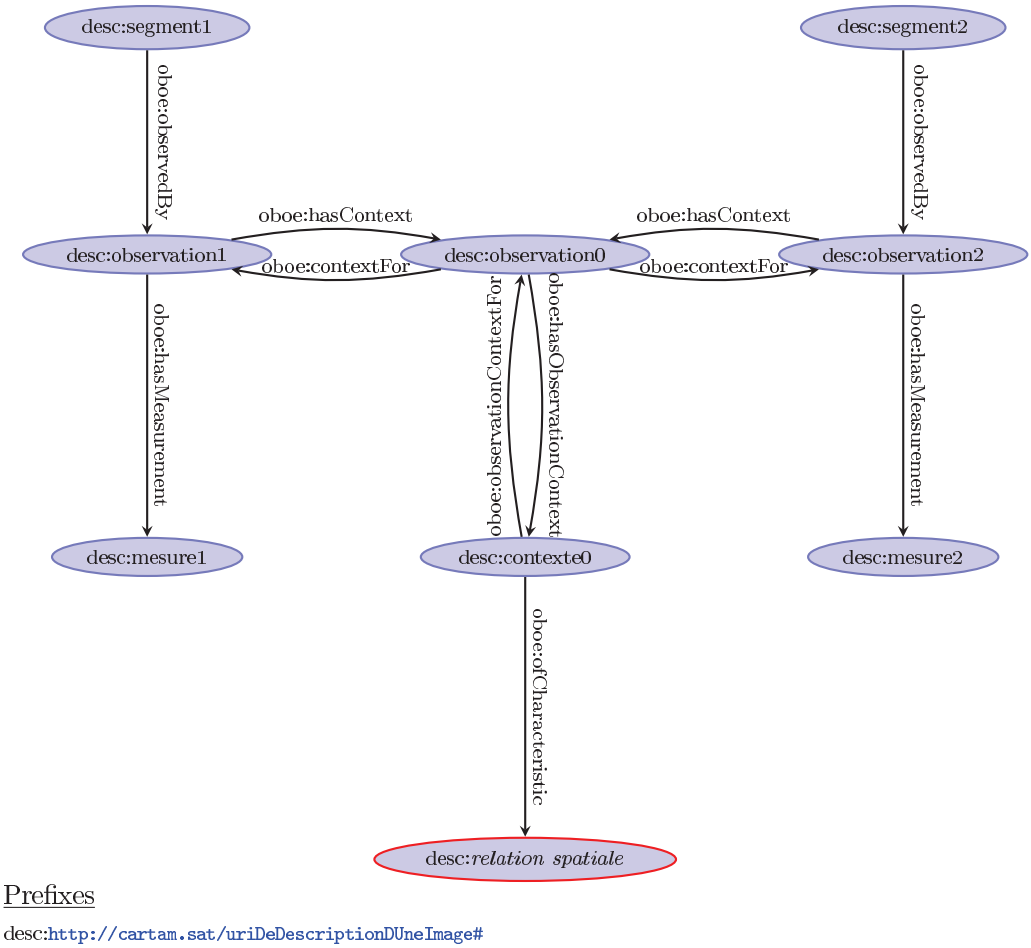


FIGURE B.1 – Première possibilité d'utilisation des relation spatiales réifiées d'OBOE. L'instance de la relation spatiale est entourée de rouge.

L'extension spatiale offerte par l'ontologie OBOE permet théoriquement de réifier les relations spatiales. Cette option méritait d'être explorée, ne serait-ce que parce que la réification d'une relation permet de la décrire, de la caractériser. Les trois concepts relatifs aux relations spatiales d'OBOE sont : *Adjacent*, *Overlaps* et *Within*. Ces concepts sont des spécialisations du concept de caractéristique et comme tels doivent se rattacher à une mesure, ou plus exactement à un *contexte*

puisqu'elle est relative à plusieurs entités.

Une première possibilité illustrée par la FIGURE B.1 est symétrique. La fonctionnalité de la propriété *observationContextFor* est respectée. Mais la symétrie qui convient si les segments *segment1* et *segment2* sont adjacents, n'est pas appropriée, par exemple, si l'un est contenu dans l'autre.

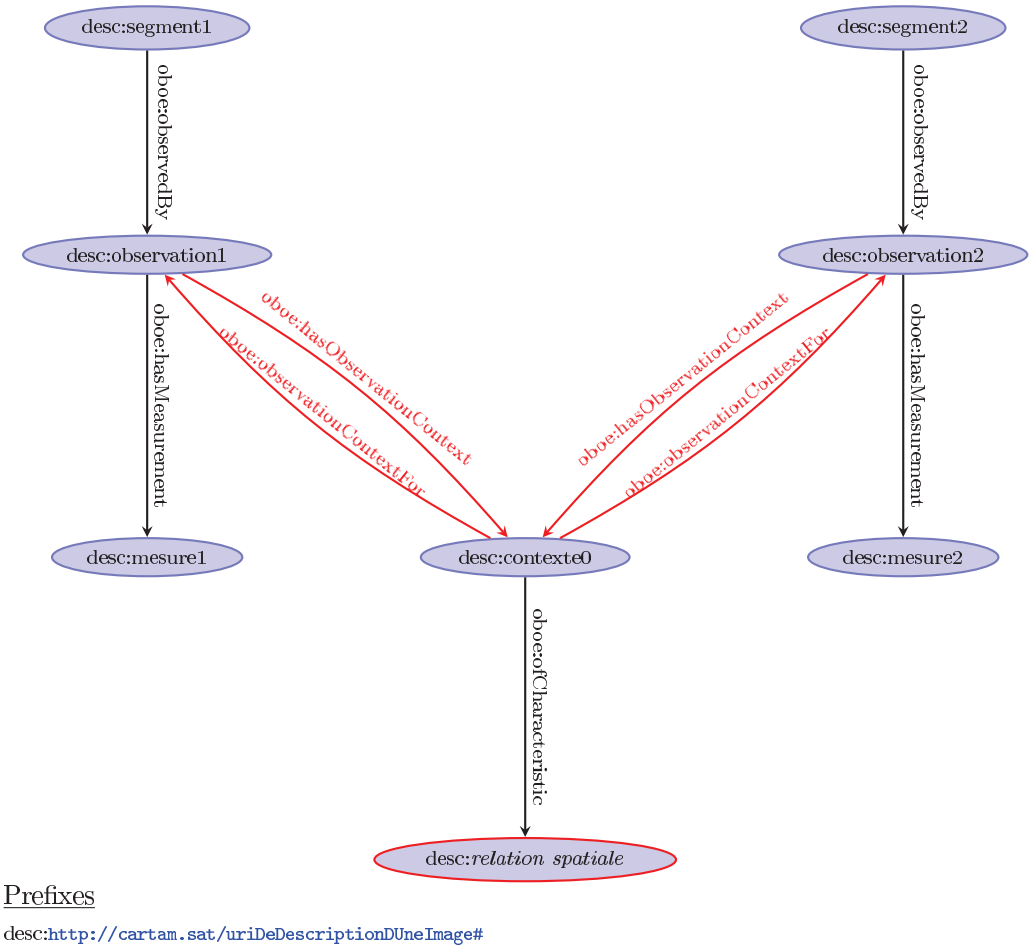


FIGURE B.2 – Deuxième possibilité d'utilisation des relation spatiales réifiées d'OBOE. L'instance de la relation spatiale est entourée de rouge. Les propriétés incohérentes sont également en rouge.

Une deuxième possibilité, qui se passe d'observation dédiée au contexte (FIGURE B.2), conserve la question de la symétrie qui n'est pas adaptée à toutes les relations spatiales et pose en plus un problème de cohérence de la base de connaissances en violant la fonctionnalité de la propriété *observationContextFor*. Cette violation qui pourrait être évitée en apparence en utilisant la propriété *hasContextObservation* serait fatalement déduite par inférence, puisque *observationContextFor* en est la propriété inverse.

OBOE propose également d'autres propriétés permettant de relier des contextes et des observations. Il s'agit de *hasContextObservation* (qui a pour domaine un contexte) et de *contextObservationFor* (qui a pour domaine une observation). Ces propriétés sont respectivement fonctionnelle et fonctionnelle inverse... mais elles ne sont pas explicitement définies comme inverses l'une de l'autre, ce qui permet d'éviter la structure symétrique de la FIGURE B.2. Cette possibilité illustrée par la FIGURE B.3 permet de d'éviter toute violation de fonctionnalité ainsi que les inconvénients d'une structure symétrique. En effet, si l'on décrit dans un sens la relation spatiale qui lie le premier segment au deuxième (le premier est inclus dans le deuxième, par une instance de *Within* par exemple), cela ne signifie plus, par absence de symétrie, que le deuxième est inclus dans le premier. Si l'on veut caractériser la relation qui lie le deuxième segment au premier, il faut alors instancier un second contexte et une seconde relation.

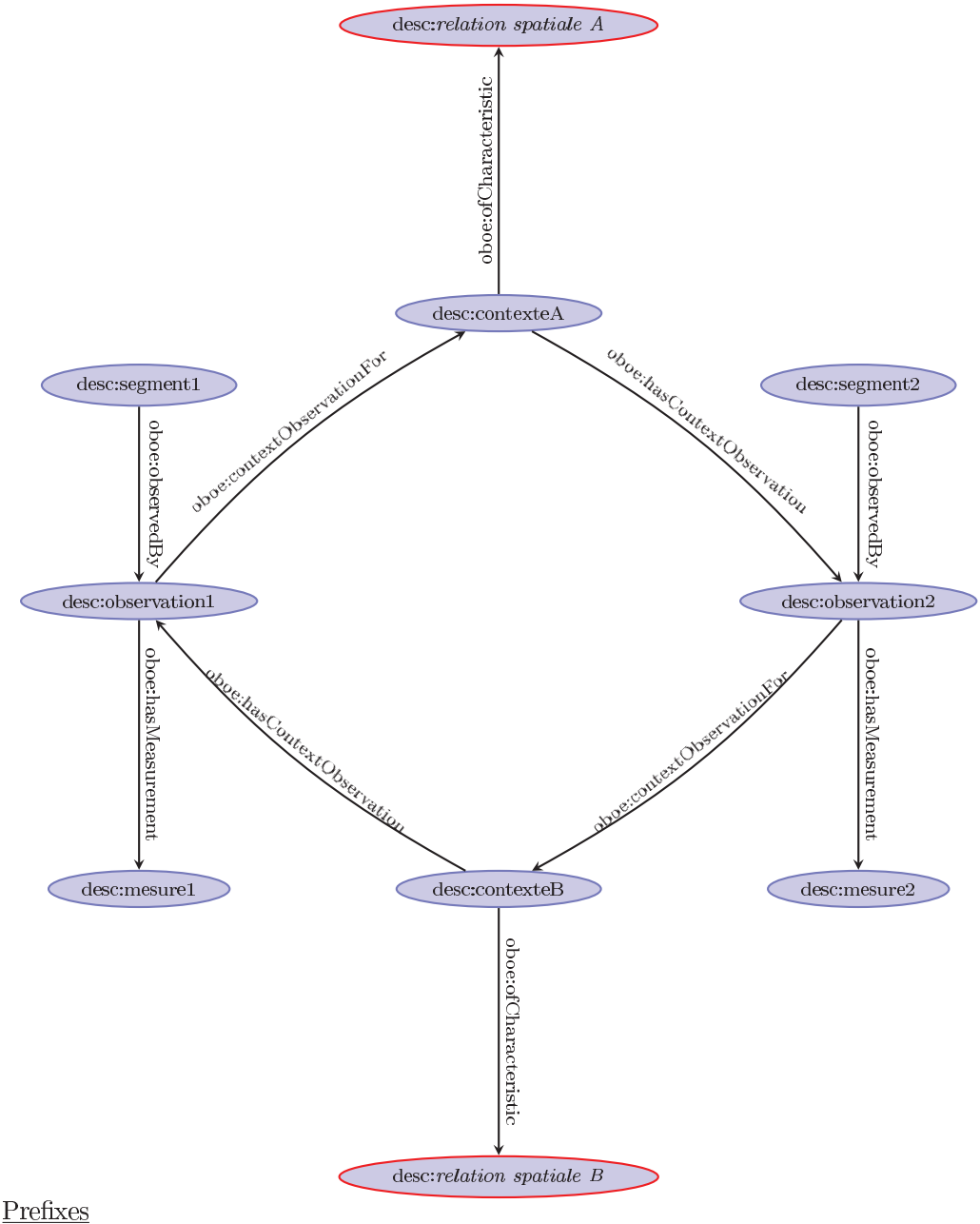


Figure B.3 – Troisième possibilité d'utilisation des relation spatiales réifiées d'OBOE. Les instances de la relation spatiale sont entourées de rouge.

Si la relation spatiale en question est l'adjacence, cela peut conduite à relier deux segments par deux instances du concept relatif à la relation d'adjacence... On pourrait éviter cette lourdeur supplémentaire en partageant l'instance de la relation spatiale entre les deux contextes, dans le cas des relations spatiales symétriques (Figure B.4).

Quoi qu'il en soit, la réification des relations spatiales proposée par OBOE ne semble reposer sur aucune norme, standard ou recommandation. Son utilisation est complexe. La duplication des propriétés permettant de relier les contextes aux observations indique probablement que l'implémentation d'OBOE a choisi de proposer aux utilisateurs plusieurs possibilités permettant d'éviter les problèmes de cohérence posés par la fonctionnalité.

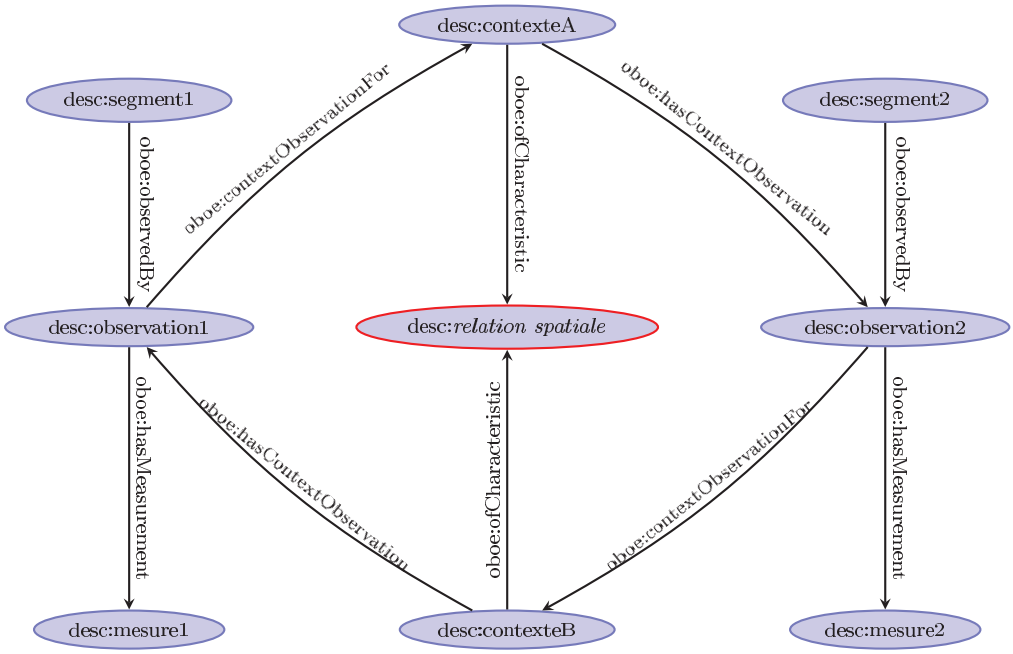


FIGURE B.4 – Troisième possibilité d'utilisation des relation spatiales réifiées d'OBOE déclinée pour les relations spatiales symétriques. L'instance de la relation spatiale est entourée de rouge.

Table des matières

Sommaire **4**

Remerciements **7**

Préambule **9**

1 Introduction **11**
1.1 La raison du sujet . . . 12
1.2 Le cadre institutionnel de la thèse . . . 12
1.2.1 L'Institut de Recherche pour le Développement . . . 12
1.2.2 L'Espace au service du Développement . . . 13
1.2.3 Systèmes d'Information et de Connaissance . . . 13
1.3 Cartographie dynamique des Territoires Amazoniens : des satellites aux acteurs . . . 14
1.3.1 Du satellite à l'image . . . 14
1.3.2 De l'image à la connaissance . . . 14
1.3.3 De la connaissance aux acteurs . . . 14
1.4 Objectifs de la thèse . . . 14
1.5 Plan du mémoire . . . 14

Conventions de notation **17**

I Contexte et problématique **19**

2 Généralités et spécificités des images en télédétection **21**
2.1 Introduction aux images numériques . . . 22
2.1.1 L'étymologie de l'image . . . 22
2.1.2 Généralités sur les images numériques . . . 23
2.2 La télédétection . . . 25
2.2.1 Processus général . . . 25
2.2.2 Lumière solaire et couleurs . . . 26
2.2.3 Télédétection spatiale . . . 28
2.2.4 Réflectance et signature spectrale . . . 35
2.2.5 Les niveaux de prétraitement des images SPOT . . . 43
2.2.6 Les résolutions d'une image satellitaire . . . 44
2.2.7 La télédétection par les hyperfréquences . . . 45
2.3 Enregistrement, traitement et analyse de l'information des images de télédétection 45
2.3.1 Nature de l'image de télédétection . . . 45
2.3.2 Format des images . . . 46

2.3.3 Mise en évidence visuelle par rehaussement et filtrage, traits des images . . 46
2.3.4 Opérations sur les images 49
2.3.5 Notion de texture 49
2.3.6 Techniques de classification appliquées aux images 52
2.3.7 Segmentation 53

3 Formalisation et conceptualisation de la connaissance 61
3.1 Bref historique de la conceptualisation et de la représentation des connaissances . . 62
3.1.1 Origine de l'ontologie 62
3.1.2 Les ontologies en informatique 63
3.2 Structuration et description des données 63
3.2.1 XML : l'ancêtre du web sémantique 63
3.2.2 RDF : tout est description 64
3.2.3 RDFS : la conceptualisation des ressources 65
3.3 Les logiques de description 67
3.3.1 Définitions élémentaires 67
3.3.2 Niveau assertionnel 70
3.3.3 Niveau terminologique 70
3.3.4 Notion d'interprétation 71
3.3.5 Les constructeurs enrichissent l'expressivité d'un langage 72
3.3.6 OWL : le web et l'expressivité des logiques de description 74
3.3.7 Complexité des logiques de description 75
3.3.8 Raisonneurs 75
3.4 Représentation de la connaissance et enrichissement de l'information 77

4 Les ontologies en imagerie, télédétection et données spatiales 79
4.1 La sémantique des pixels 80
4.1.1 Sémantique objective et sémantique subjective 80
4.1.2 Sémantique physique et sémantique de télédétection, préclassification de Baraldi et al. (2006) 80
4.2 Analyse d'image orientée objet 82
4.2.1 Le rôle de la segmentation dans la création des objets 82
4.2.2 Objet géographique et objet de l'image 83
4.2.3 Notion de point de vue 83
4.3 Segmentation et sémantique 85
4.3.1 Impact de la segmentation sur la sémantique des segments 85
4.3.2 Segmentation guidée par de la connaissance 86
4.4 Les ontologies pour l'annotation d'images 86
4.5 Formalisation de la connaissance experte 86
4.5.1 Interprétation sémantique par un algorithme dédié 86
4.5.2 Ontologies et apprentissage 87
4.5.3 Les ontologies en imagerie médicale et autres domaines 87
4.5.4 Conclusion sur les différentes applications 88
4.6 Raisonnement 88
4.6.1 L'aboutissement de la connaissance formalisée 88
4.6.2 Raisonner sur le contenu d'une image 88
4.7 Structuration des données spatiales 89
4.8 Réflexions sur la dimension temporelle 90

II Proposition d'approche 93

5 Principe général de l'approche 95
5.1 Conceptualisation de référence 96
5.1.1 Ontologies de la conceptualisation de référence 96
5.1.2 Concurrence des conceptualisations de référence 100

5.1.3 Objet de la conceptualisation de référence . . . 102
5.2 Connaissance contextuelle . . . 103
5.3 Peuplement par la description de l'image . . . 105
5.4 Interprétation sémantique de l'image . . . 106
5.4.1 Opération d'union sur les bases de connaissance . . . 106
5.4.2 Analogie avec la règle de trois et le produit en croix . . . 106

6 Planification de la mise en œuvre 109
6.1 Données en entrée . . . 110
6.1.1 Quelles données ? . . . 110
6.1.2 Quelle bonne segmentation ? . . . 110
6.2 Que faire de la connaissance en sortie de la segmentation ? . . . 113
6.2.1 Décrire la connaissance utilisée pour la segmentation . . . 113
6.2.2 Ignorer la connaissance utilisée pour la segmentation . . . 114
6.3 Intégration et modularité de la connaissance . . . 116
6.3.1 La source de la connaissance experte . . . 116
6.3.2 La place de la connaissance par rapport aux traitements . . . 117
6.4 Réflexions sur la modularisation de la connaissance . . . 121
6.4.1 Pourquoi modulariser la connaissance ? . . . 121
6.4.2 Comment modulariser la connaissance ? . . . 122
6.5 Réflexions sur la représentation des relations spatiales . . . 122
6.5.1 Représentation des relations topologiques . . . 122
6.5.2 Interprétation de l'adjacence dans une image matricielle avec RCC8. . . . 123

7 Déclinaisons du principe et implémentations 125
7.1 Application associée à l'analyse orientée objet . . . 126
7.1.1 Analyse orientée objet avec *Orfeo Toolbox* . . . 126
7.1.2 Description de l'image segmentée . . . 127
7.1.3 Exploitation de la description de l'image segmentée . . . 128
7.1.4 Proposition d'évaluation . . . 131
7.1.5 Conclusion de la validation de l'approche . . . 134
7.2 Intégration des relations spatiales . . . 135
7.2.1 Vers la définition de nouveaux concepts . . . 135
7.2.2 Proposition d'évaluation . . . 136
7.2.3 Limites de l'utilisation de RCC8 . . . 138
7.2.4 Potentiel de la parallélisation . . . 140
7.2.5 Influence de la segmentation et des caractéristiques de l'image . . . 140
7.3 Adaptation à la description des pixels . . . 141
7.3.1 Une déclinaison *a priori* difficile . . . 141
7.3.2 Contourner la complexité des logiques de description par une hypothèses de travail . . . 142
7.3.3 Vers une segmentation ontologique ? . . . 143
7.4 Classification séparée des objets . . . 145
7.4.1 Comparaison des classifications globale et séparée . . . 145
7.4.2 Produits du raisonnement : du masque sémantique à la carte interprétée . . . 145
7.5 Réflexions . . . 148
7.5.1 Valeur ajoutée de l'approche . . . 148
7.5.2 Les enjeux du passage à l'échelle . . . 148
7.5.3 Vers une approche ontologique hybride pixels/objets ? . . . 148
7.6 L'hypothèse du monde ouvert . . . 149
7.6.1 Comment définir les concepts ? . . . 149
7.6.2 Comment décrire un fait ? . . . 152
7.6.3 Comment regarder les résultats ? . . . 152

III Conclusion **155**

8 Conclusion et ouverture **157**
8.1 Bilan . . . 158
8.1.1 Résumé structurel . . . 158
8.1.2 Développement de l'approche sur trois axes . . . 158
8.2 Remarques relatives au déroulement des travaux . . . 159
8.2.1 Risques inhérents au positionnement . . . 159
8.2.2 Communiquer et partager la connaissance . . . 160
8.2.3 Discussion des performances . . . 160
8.2.4 Difficulté de l'interaction avec les experts . . . 161
8.3 Perspectives . . . 161
8.3.1 En aval : les thématiques de terrain . . . 161
8.3.2 En amont : les prétraitements . . . 163
8.3.3 En profondeur : étiquetage des traitements . . . 164
8.3.4 En largeur : réflexions sur l'identification d'individus composés . . . 164
8.3.5 Dans l'immédiat : enrichir les descriptions des concepts de télédétection . . . 164

IV Annexes **167**

A Quelques indices utilisés en traitement d'image **169**
A.1 Indices spectraux . . . 169
A.2 Indices de texture . . . 170

B Essais d'utilisation des relations spatiales réifiées d'OBOE **171**

Table des matières **175**

Liste des définitions et des remarques **179**

Table des figures **182**

Liste des tableaux **188**

Bibliographie **189**

Liste des définitions et des remarques

1 Définition (Système d'information, d'après Reix (2004)) 13
2 Définition (Système d'information géographique) 13

3 Définition (Maillage, d'après Bonnet (2010)) 24
4 Définition (Pavage, d'après Manzanera (2011); Bonnet (2010)) 24
5 Définition (Pixel, d'après Manzanera (2011)) 24
6 Définition (Définition) 24
7 Définition (Résolution spatiale) 24
8 Définition (Résolution tonale, d'après Manzanera (2011)) 24
1 Remarque (Niveaux de gris, d'après Manzanera (2011)) 24
9 Définition (Dynamique d'image, d'après Manzanera (2011)) 25
1 Corollaire (Dynamique d'image (Manzanera, 2011)) 25
10 Définition (Télédétection) 25
11 Définition (Photointerprétation) 25
2 Remarque (Télédétection et photointerprétation) 25
12 Définition (Couleur, d'après Bonnet (2005)) 27
13 Définition (Image monospectrale, canal) 28
14 Définition (Image spectrale) 28
15 Définition (Image multispectrale) 29
16 Définition (Image panchromatique) 29
3 Remarque (Résolution spatiale en panchromatique et en multispectral) 29
4 Remarque (Expression courante de la résolution spatiale) 29
17 Définition (Résolution temporelle) 30
5 Remarque (Résolution temporelle) 30
18 Définition (Bande spectrale) 30
6 Remarque (Bande spectrale, canal) 31
19 Définition (Photogrammétrie) 32
20 Définition (Orthorectification, ou redressement) 34
21 Définition (Géoréférencement) 35
22 Définition (Métadonnée) 35
7 Remarque (Métadonnée, d'après le *World Wide Web Consortium*) 35
8 Remarque (Métadonnée géographique, d'après la présentation de la norme ISO 19115 :2003) 35
23 Définition (Irradiance ou éclairement énergétique ($W \cdot m^{-2}$), d'après Pouget (2005)) 36
24 Définition (Luminance énergétique ($W \cdot m^{-2} \cdot sr^{-1}$), d'après Pouget (2005)) 36
25 Définition (Réflectance, d'après Pouget (2005)) 38
9 Remarque (Réflectance spectrique, d'après Bonnet (2005)) 38

26 Définition (Réflectance spectrique) . . . 38
10 Remarque (Réflectance spectrique, angle d'incidence, angle d'observation et couleur (Bonnet, 2005)) . . . 38
11 Remarque (Réflectance et réflectivité) . . . 38
27 Définition (Signature spectrale) . . . 39
12 Remarque . . . 39
28 Définition (Vérité terrain) . . . 40
29 Définition (Calibration radiométrique (absolue), d'après Bonn (1994)) . . . 41
13 Remarque (Calibration radiométrique) . . . 41
14 Remarque (Calibration radiométrique relative, d'après Bonn (1994)) . . . 41
30 Définition (Résolution radiométrique) . . . 44
31 Définition (Résolution spectrale) . . . 44
15 Remarque (Relation entre les résolutions spatiale, radiométrique et spectrale) . . . 45
32 Définition (Histogramme) . . . 47
33 Définition (Corrélogramme, d'après Xiaoguang (2002)) . . . 48
34 Définition (Indice spectral) . . . 49
35 Définition (Image pseudospectrale) . . . 49
36 Définition (Matrice de co-occurrence en niveaux de gris, d'après Haralick et al. (1973)) 50
16 Remarque (Résolution tonale et taille de la GLCM) . . . 51
37 Définition (Image texturale) . . . 51
17 Remarque . . . 51
38 Définition (Classification) . . . 52
39 Définition (Segmentation, d'après Castilla and Hay (2008)) . . . 53
18 Remarque (Segmentation) . . . 53
40 Définition (Image de segmentation) . . . 53
19 Remarque (Image de segmentation et vectorisation) . . . 53
41 Définition (Segment) . . . 53
42 Définition (Objet de l'image, d'après Castilla and Hay (2008)) . . . 53
20 Remarque (Bonne segmentation) . . . 53
43 Définition (Entité de l'image) . . . 54
44 Définition (Entité (ou objet) de terrain (ou géographique)) . . . 54
45 Définition (Entité géoréférencée) . . . 54
21 Remarque (Segmentation et classification des pixels) . . . 54
Pour résumer (Image et information) . . . 59

46 Définition (Ontologie, d'après Gruber (1993)) . . . 63
47 Définition (Ressource) . . . 64
22 Remarque . . . 64
48 Définition (Logique de description, d'après Baader and Nutt (2003); Baader et al. (2007)) . . . 67
49 Définition (Décidabilité (algorithmique), d'après Bournez et al. (2010)) . . . 67
50 Définition (Individu) . . . 68
51 Définition (Concept, d'après Nardi and Brachman (2003)) . . . 68
23 Remarque (Notation des concepts) . . . 68
52 Définition (Concepts atomique, composé, primitif et défini, d'après Baader and Nutt (2003)) . . . 69
53 Définition (Relation binaire, d'après Deschamps et al. (1999)) . . . 69
54 Définition (Rôle, d'après Nardi and Brachman (2003)) . . . 69
24 Remarque (Notation des rôles) . . . 69
25 Remarque (Axiomes assertionnels (ou assertions) d'appartenance et de relation) . . . 70
Pour résumer (Abox) . . . 70
26 Remarque (D'après Nardi and Brachman (2003)) . . . 70
27 Remarque (Concept et expression symbolique) . . . 71
Pour résumer (TBox) . . . 71
55 Définition (Base de connaissances, d'après Baader et al. (2007)) . . . 71
56 Définition (Modèle, d'après Baader et al. (2007)) . . . 72

57 Définition (Consistance, d'après Baader et al. (2007)) 72
58 Définition (Satisfaisabilité, d'après Baader and Nutt (2003)) 72
59 Définition (Expressivité) 72
60 Définition (Domaine concret, d'après Baader and Sattler (1998)) 74
61 Définition (Classification d'une base de connaissances, d'après Bock et al. (2008)) . 75
62 Définition (Réalisation d'une base de connaissances, d'après Bock et al. (2008)) . . 75
28 Remarque (Réalisation et classification) 76
63 Définition (Hypothèse du monde ouvert, d'après Hustadt (1994)) 77
Pour résumer (Apport de la représentation des connaissances) 78

64 Définition (Fossé sémantique, d'après Smeulders et al. (2000)) 80
65 Définition (Fossé sensoriel, d'après Smeulders et al. (2000); Bissol (2006)) 84
Pour résumer 91

66 Définition (Conceptualisation de référence, adaptée d'après Falomir et al. (2011)) . 96
67 Définition (Contexte de validité (d'une conceptualisation ou d'une connaissance)) . 96
Pour résumer (Conceptualisation de référence) 103
68 Définition (Connaissance contextuelle, adaptée d'après Falomir et al. (2011)) . . . 103
1 Hypothèse (Hypothèse distinction des entités de l'image des entités de terrain) . . 104
Pour résumer (Connaissance contextuelle) 105
Pour résumer (Faits de l'image) 106
69 Définition (Union sur les bases de connaissances) 106

2 Hypothèse (Hypothèse de généricité de l'approche) 110
3 Hypothèse (Hypothèse de calibration de l'image en réflectance et température) . . 110
4 Hypothèse (Hypothèse de bonne segmentation) 111
5 Hypothèse (Hypothèse de définition intrinsèque des segments) 111

6 Hypothèse (Hypothèse de description des segments) 126
7 Hypothèse (Hypothèse d'interprétation par raisonnement) 126
1 Résultat (Raisonnement sur une base de connaissances contenant une description complète de l'image suite à une analyse orientée objet) 134
2 Résultat (Conclusions de l'intégration des relations topologiques spatiales à l'approche) 141
8 Hypothèse (Hypothèse de la classification des pixels indépendamment les uns de autres) 142
3 Résultat (Conclusions de l'adaptation de l'approche aux pixels) 144
4 Résultat (Conclusions de la classification séparée des objets) 147

Table des figures

1.1 Conventions de notation graphique. 17
1.2 Conventions de notation des éléments en logiques de description. 18
1.3 Conventions de notation des chaînes de traitements et descriptifs assimilés, d'après le langage formalisé par Lin (2011). 18

2.1 Principe de la télédétection. 26
2.2 Le spectre électromagnétique. 28
2.3 Le système de balayage perpendiculaire (à miroir rotatif). Dans chaque ligne, chaque pixel est enregistré successivement, au fur et à mesure de la rotation du miroir associé au capteur. Les lignes se succèdent au fur et à mesure du déplacement du satellite. Avec ce type de capteur, les pixels d'une même ligne ne sont pas enregistrés au même instant et ils correspondent donc chacun à une position différente du satellite. . . . 33
2.4 Le système de balayage parallèle (à barrette de capteurs). Dans chaque ligne, tous les pixels sont enregistrés simultanément. Les lignes se succèdent au fur et à mesure du déplacement du satellite. Avec ce type de capteur, les pixels d'une même ligne sont enregistrés au même instant et correspondent donc à une unique position du satellite. 33
2.5 Les étapes de calibration radiométrique. L'information principale est représentée dans un ovale, l'information extérieure dans un rectangle. Le CODE 2.3 donne un exemple concret de métadonnées relatives aux caractéristiques des capteurs et de l'éclairement. L'information sur les caractéristiques atmosphériques sont généralement modélisées ou acquises via d'autres instruments. Elle permet d'estimer plus précisément la réflectance « au sol », c'est-à-dire corrigée des perturbations atmosphériques. 43
2.6 Le processus d'enrichissement de l'information. L'information principale est représentée dans un ovale, l'information extérieure dans un rectangle. 43
2.7 Exemple de rehaussement d'image. En (B), un canal représenté en niveaux de gris est extrait de l'image multispectrale d'origine (A) représentée ici en fausses couleurs. Mais les données du canal en question n'exploitent pas toutes les valeur possibles (voir FIGURE 2.8). En recalculant les valeurs des pixels de manière à exploiter l'ensemble des valeurs de 0 à 255 (voir FIGURE 2.9), l'image (C) montre que l'on améliore le contraste visuel. 47
2.8 Exemple d'histogramme : la répartition des pixels en fonction de leur valeur (de 0 à 255) dans l'image (B) de la FIGURE 2.7. À l'observation des extrémités de l'histogramme, on constate qu'aucun pixel n'a de valeur proche de 0 ou de 255. À l'affichage, cela se traduit par l'absence de blanc et de noir. 48
2.9 Exemple d'histogramme : l'effet du rehaussement linéaire sur l'histogramme dans l'image (C) de la FIGURE 2.7 (à comparer avec l'histogramme en FIGURE 2.8). . . 48

2.10 Exemple de calcul d'indice (B) sur une image multispectrale de Cayenne (A). L'indice en question est le NDVI (*Normalized Difference Vegetation Index*) calculé à partir des bandes rouge et proche infrarouge (c'est-à-dire les canaux 2 et 3 précisés en TABLE 2.2). La couverture végétale (en rouge sur l'image en fausses couleurs) correspond à un fort NDVI : les zones les plus claires sur l'image (B). Les zones les plus sombres qui correspondent à de l'eau ont un NDVI très faible. Dans la zone hétérogène (urbanisée) le NDVI n'est pas négligeable. Dans l'absolu, cela pourrait correspondre à la présence de végétation chlorophyllienne en ville ou bien simplement à un pouvoir séparateur moindre du NDVI entre la végétation et l'urbanisme qu'entre la végétation et l'eau. 50
2.11 Exemple de calcul de texture (B) sur une image multispectrale de Cayenne (A). La texture en question est une *inertie* calculée sur le moyen infrarouge : les zones foncées (faible inertie) sur l'image (B) peuvent être mises en correspondance des zones hétérogènes de l'image d'origine (A). Inversement, l'inertie des zones homogènes est plus élevée. 52
2.12 Exemple de segmentation (arbitraire) d'une image multispectrale de Cayenne (A). La valeur d'un pixel de l'image en sortie de la segmentation correspond à l'identifiant du segment d'image dont il fait partie. Le résultat peut donc être rigoureusement visualisé en niveaux de gris. La coloration des segments (présentée ici) permet toutefois d'améliorer la visualisation. Les premières difficultés de l'étape de segmentation apparaissent : alors que la composante urbaine de l'image apparaît par exemple bien plus hétérogène que les régions végétalisées, les segments découpés dans ces deux espaces ne présentent pas de différence de taille caractéristique des principaux environnements qui apparaissent sur l'image : eau, végétation, urbanisme. 55
2.13 Impact du contexte général de l'image sur la segmentation au niveau local avec l'algorithme de déplacement de la moyenne. L'algorithme est appliqué deux fois, avec des paramètres identiques, mais sur deux extraits différents de l'image (I) partageant une étendue commune. Une observation locale de la zone (L) située dans l'étendue commune, montre que le découpage de l'image diffère, pour une même zone (avec un algorithme et des paramètres identiques) en fonction de l'extrait d'image considéré. Cet exemple illustre la difficulté que peut poser la segmentation dans l'association des objets de l'image avec des caractéristiques qui leur seraient propres. En d'autres termes : *comment définir correctement un objet de l'image dont les contours dépendent de pixels représentant des régions situées à des dizaines de kilomètres de distance ?* 56
2.14 Comparaison visuelle de l'effet de différents algorithmes de segmentation appliqués à un extrait d'image satellitaire (A). Les paramètres des algorithmes de partage des eaux (B) de déplacement de la moyenne (C) n'ont pas été ajustés. Les régions produites sont difficiles à rapprocher des grands ensembles (eau, végétation, urbain) qui apparaissent sur l'image. Une méthode de segmentation par seuillage appliquée à la suite d'une préclassification des pixels de l'image d'origine (D) permet au contraire d'opérer une certaine interprétation sémantique de nombreuses zones, sans qu'il ait été nécessaire d'ajuster de quelconques paramètres. De plus, la taille des segments en (D) est beaucoup plus hétérogène qu'en (C) et surtout qu'en (B). Certaines méthodes de segmentation sont donc plus adaptées que d'autres à la délimitation de segments de tailles hétérogènes. 57
2.15 Du pixel au contexte des objets dans l'image. 58

3.1 Un exemple de graphe RDF. 64
3.2 Une illustration du mécanisme de réification dans RDF. Un triplet peut être vu comme une ressource et jouer le rôle d'un objet ou d'un sujet dans un des triplets. Les prédicats, représentés par des flèches, sont de trois sortes. Ceux de couleur noire indiquent que le triplet réifié, comme ressource, peut jouer le rôle d'un sujet ou d'un objet dans d'autres triplets. Le prédicat de couleur bleue représente explicitement le typage de la ressource constituée par le triplet. Enfin, les prédicats de couleur rouge indiquent le sujet, le prédicat et l'objet du triplet réifié. 67

3.3 La représentation du monde par les logiques de description. Des entités du monde telles que des pixels (à gauche) sont représentées par des individus (à droite) généralement dotés d'un identifiant (#i). Les individus peuvent appartenir à des ensembles : ici, le concept de *Pixel*. 68
3.4 Les relations entre individus sont des couples d'individus. Ici, les individus représentent quelques pixels de différentes sortes. Une relation d'adjacence lie un certain nombre d'individus entre eux. L'individu #1, par exemple, est adjacent au #2. L'ensemble de ces couples définit le rôle *Adjacent* sur $\mathcal{P}ixel \times \mathcal{P}ixel$. 69
3.5 Une fonction d'interprétation $.^{\mathcal{I}}$ associe à chaque individu, un élément du domaine d'interprétation $\Delta^{\mathcal{I}}$ (ici, des pixels). Toutefois, dans cette illustration, l'interprétation $(.^{\mathcal{I}}, \Delta^{\mathcal{I}})$ ne constitue pas un modèle : on a en effet $(PixelEau(\#1) \wedge \neg PixelEau^{\mathcal{I}}(\#1^{\mathcal{I}}))$. 72
3.6 Le principe de classification d'une $\mathcal{T}Box$ par un raisonneur. 77
3.7 Le principe de réalisation d'une $\mathcal{A}Box$ par un raisonneur. 78

4.1 Sémantique objective d'une entité. Le fait de rattacher une entité à un concept lui confère une sémantique. 81
4.2 Sémantiques subjectives d'une entité. Plus on considère un niveau subjectif de sémantique d'une entité, plus celle-ci est susceptible de séparer les observateurs, même si ceux-ci sont d'accord sur une sémantique plus objective. Ici, deux acteurs s'accordent sur le concept de pixel. Mais ils diffèrent par leur interprétation de ses caractéristiques, ce qui les conduit à lui attribuer des sémantiques subjectives différentes. 81
4.3 Une illustration de la notion de point de vue, d'après Mariño Drews (1993). Celle-ci découle principalement de la différence d'objectifs entre les observateurs (A) et (B) qui les pousse à ne « voir » que les caractéristiques qui les intéressent. 84
4.4 Une illustration de l'acception étendue de la notion de point de vue. On y retrouve le *fossé sémantique* entre les données numériques et leur interprétation. On peut également y placer le *fossé sensoriel*. Bien qu'ils s'intéressent à la même entité de terrain (id), les observateurs (A) et (B) décrivent deux entités différentes : respectivement l'entité géographique de terrain (id) et l'entité de l'image (id') qui la représente. . . 85

5.1 Un extrait de la conceptualisation de référence : les principaux concepts relatifs à la structure de l'image utilisés par le logiciel d'analyse d'images de télédétection dans le cadre de la thèse. 97
5.2 Un extrait de la conceptualisation de référence : les principaux concepts relatifs aux traitements de l'image utilisés par le logiciel d'analyse d'images (*Orfeo Toolbox*) de télédétection dans le cadre de la thèse. 99
5.3 Un extrait de l'ontologie OBOE. 100
5.4 Un extrait de l'ontologie de conceptualisation de référence de l'expert. Par opposition, la FIGURE 5.1 et la FIGURE 5.2 étaient relatives à la conception de l'image du point de vue de la brique logicielle. Cette différence apparaît dans les différences d'espaces de noms utilisés dans ces deux dernières figures, et la représentation ci-dessus. 101
5.5 La conceptualisation de référence est la base de la description du contenu de l'image par instanciation des concepts de référence et la base de la description de la connaissance humaine par spécialisation des concepts de référence. 102
5.6 Construction type simplifiée d'un concept de la connaissance contextuelle. À droite de l'axiome d'équivalence, on trouve des concepts atomiques (en rouge), des rôles (en orangé) et des propriétés sur des domaines concrets (en jaune). Tous ces éléments sont issus de la conceptualisation de référence et combinés à l'aide de constructeurs logiques (en vert). À gauche de l'axiome d'équivalence, le concept atomique défini dans la connaissance contextuelle. 103
5.7 Description de la connaissance contextuelle à partir de la conceptualisation de référence. Il s'agit principalement de spécialiser les concepts de la conceptualisation de référence pour chaque concept de télédétection que l'on souhaite représenter. . . . 105

5.8 Description des faits de l'image à partir de la conceptualisation de référence. Il s'agit principalement d'instancier les concepts de la conceptualisation de référence pour chaque entité de l'image que l'on souhaite représenter. 106
5.9 Combinaison des différentes descriptions nécessaires à l'interprétation sémantique des faits de l'image. 107

6.1 L'extrait (825 × 666 pixels) de l'image Landsat 5 utilisée dans le développement d'une approche intégrée, dans une combinaison ici des longueurs d'onde visibles. . 111
6.2 Le résultat coloré de la préclassification de BARALDI appliquée à l'image présentée en FIGURE 6.1. Ici, chaque couleur représente une préclasse particulière. 112
6.3 Le résultat coloré de la segmentation basée sur la préclassification de BARALDI présentée en FIGURE 6.2. Ici, chaque région colorée représente un segment d'image. 112
6.4 Différentes pistes d'utilisation des ontologies dans les images satellitaires. En privilégiant une description ontologique quantitative (en rouge) par rapport à une description qualitative (en bleu) on rapproche d'autant la connaissance du monde de l'image. La piste qualitative offre deux possibilités : décrire simplement une interprétation classique, ou bien l'affiner par raisonnement automatique. De la même manière, une approche quantitative permet soit de décrire simplement le contenu de l'image, soit de raisonner sur cette description. C'est cette dernière piste que nous avons privilégiée. Des considérations plus techniques permettent ensuite d'envisager, soit un raisonnement assez fortement intégré au traitement d'image, soit un raisonnement séparé. 115
6.5 Schéma de principe de la validation de l'approche théorique. Le processus part d'une image satellite et aboutit à sa description sémantique. La connaissance structurée (en vert) intervient en aval du traitement d'image. La description conceptualisée en triplets se réfère à la conceptualisation de référence, mais n'utilise pas le fichier OWL dans lequel elle est définie (d'où le lien grisé). 119
6.6 Schéma de principe d'intégration du raisonnement (voir la FIGURE 6.5 pour la partie amont). 120
6.7 Nouveau schéma de principe d'intégration du raisonnement (voir la FIGURE 6.5 pour la partie amont). 121
6.8 Application de la relation RCC8 *Externally Connected* aux images matricielles. En considérant la partition de l'image matricielle proposée en haut, on désire savoir si les segments 4 et 5 (qui ne sont pas adjacent au sens commun de ce mot) sont liés par la relation RCC8 *Externally Connected*. Il faut pour cela les isoler comme (présenté au milieu), avant d'appliquer une dilatation d'un pixel correspondant à la définition de leurs frontières (représentation du bas). On constate alors que l'intersection des frontières n'est pas vide, ce qui implique que les segments sont *Externally Connected*. 123

7.1 Rappel de la partie amont (analyse de l'image) des chaînes de traitements exposées en FIGURE 6.5, FIGURE 6.6 et FIGURE 6.7. 127
7.2 Schéma simplifié de la description ontologique de l'image avec *Soprano* (version *non-intégrée*). La description des faits est similaire en ce qui concerne l'alimentation d'un raisonneur (version *intégrée*). Au lieu d'alimenter un modèle RDF, les triplets prennent la forme d'assertions qui alimentent la base de connaissances du raisonneur. 128
7.3 Le segment d'image et ses interprétations sémantiques. 129
7.4 L'éditeur d'ontologies *Protégé*. Le cadre inférieur droit montre un résultat inféré : l'individu représentant le segment d'image #80 a été classifié dans le concept de segment végétal, sélectionné dans le cadre supérieur gauche. 131
7.5 Visualisation, en blanc, des segments de l'extrait (825 × 666 pixels) de l'image Landsat 5, dotés d'une sémantique végétale. Les segments en noir n'ont pas été classifiés dans cette sémantique. 132
7.6 Sémantique de végétation, obtenue par raisonnement sur des objets (A) et par seuillage de pixels (B). Les entités végétales sont en blanc. 133

7.7 Sémantique de surfaces minérales ou construites, obtenue par raisonnement sur des objets (A) et par seuillage de pixels (B). Les entités minérales ou construites sont en blanc. 133
7.8 Sémantique aquatique, obtenue par raisonnement sur des objets (A) et par seuillage de pixels (B). Les surfaces en eau sont en blanc. 134
7.9 Définition d'une sémantique de segment minéral *à proximité* d'un segment d'eau. . 135
7.10 Utilisation de la connaissance sur une relation spatiale dans un extrait d'image Landsat à proximité de Santarém (A). En (B), une interprétation automatique sans relations spatiales. Les segments d'eau sont représentés en bleu, les segments végétaux en vert et les segments minéraux ou construits en gris. Cette première interprétation permet de détecter des combinaisons de concepts de télédétection (végétation et minéral en jaune, eau et minéral en bleu clair) dont l'interprétation relève des experts du domaine. D'autres concepts peuvent être construits à l'aide de relations spatiales, par exemple pour la détection des rives (C). Les segments qui s'y rapportent sont représentés en orange. 136
7.11 Sémantique de plage ou de rive obtenue par raisonnement sur des objets (A) et par l'expertise manuelle de deux experts en télédétection (B) et (C) (respectivement Mme Frédérique SEYLER et M. Jean-François FAURE). Les surfaces de rives sont en blanc. 137
7.12 Artefacts liés à l'implémentation du calcul des relations spatiales RCC8 dans les images matricielles. Deux segments de plage (en orange) ne sont adjacents à aucun segment d'eau. 138
7.13 Utilisation de la connaissance sur une relation spatiale dans un extrait d'image Landsat à proximité de Santarém (A). En (B), une interprétation automatique sans relations spatiales. Dans l'image (C), les zones de végétation *ExternallyConnected* à des zones d'eau sont représentées en violet. 139
7.14 Évolution du temps d'exécution du raisonnement sur les pixels en fonction de leur nombre. Cette courbe est basée sur les données de la TABLE 7.8. 142
7.15 Évolution du temps d'exécution du raisonnement sur les pixels en fonction de leur nombre. La courbe rouge, quasi-linéaire montre le gain de temps d'exécution par rapport au raisonnement sur une base de connaissances globale des faits de l'image (rappelée en bleu). Cette courbe est basée sur les données de la TABLE 7.8. 144
7.16 Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique végétale. Les segments en noir n'ont pas été classifiés dans cette sémantique. 146
7.17 Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique minérale ou construite. Les segments en noir n'ont pas été classifiés dans cette sémantique. 146
7.18 Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique aquatique. Les segments en noir n'ont pas été classifiés dans cette sémantique. 147
7.19 Composition colorée des résultats d'extraction des sémantiques végétale (vert), minérale (gris) et aquatique (bleu foncé). D'autres couleurs apparaissent indiquant que certains segments ont été classifiés dans plusieurs classes. Le jaune indique des segments végétaux et minéraux, le bleu clair des segments minéraux et aquatiques. D'un point de vue expert, cela ne signifie pas que ces entités superposent toutes les sémantiques qui leurs sont affectées. Peut-être cela illustre-t-il simplement l'existence d'autres sémantiques comme par exemple des surfaces cultivées, des eaux peu profondes etc. 147
7.20 Extraction, en blanc, des segments de l'extrait (825×666 pixels) de l'image Landsat 5 dotés d'une sémantique végétale. Les segments en noir n'ont pas été classifiés dans cette sémantique. 150

7.21 Résultat auquel on pourrait s'attendre en construisant un concept de *Segment non végétalisé* à partir du concept de *Segment végétalisé* en utilisant le constructeur de négation sur les concepts. En réalité, cette projection de résultat supposerait que tout segment qui n'a pas été classifié comme relatif à de la *végétation* soit relatif à de la *non-végétation*. Cela n'est absolument pas vrai dans le cadre de l'hypothèse du monde ouvert. En revanche, ce type de résultat qui pourrait être obtenu par le parcours d'un arbre de décision. 151
7.22 Résultat obtenu si l'on définit un concept de *Segment non végétalisé* à partir du concept de *Segment végétalisé* en utilisant le constructeur de négation sur les concepts. Aucun segment ne répond à l'appel, car du fait de la négation, il ne suffit pas qu'il n'ait pas les caractéristiques d'un *Segment végétalisé*. Il aurait fallu que la description de l'image indiquât de manière explicite ou implicite que les segments en question n'en sont pas dotés. Dans le cas contraire, ils sont toujours susceptibles de l'être, éventuellement dans une autre description. 151

8.1 Extension envisagée de l'approche en aval de l'interprétation de l'image, vers des cartes thématiques. La partie explorée au cours de la thèse (en noir) et celle explorée par Arvor et al. (2013a) (en vert) sont susceptibles d'être reliées (en rouge) par la formalisation dans une ontologie de la connaissance nécessaire pour passer d'une interprétation des entités de l'image en classes d'occupation du sol à des représentations thématiques. 162
8.2 Une possibilité de liaison entre les descriptions des objets de l'image et celles des objets de terrain pourrait consister à considérer que les objets de l'image sont le fruit d'observations d'objets de terrain en utilisant l'ontologie OBOE. Alors que l'observation d'un objet de l'image peut être caractérisée par un traitement, l'observation d'un objet de terrain pourrait être caractérisée par le capteur utilisé. 163
8.3 Récapitulatif de quelques perspectives en amont, représenté sous la forme d'une chaîne de traitements. La thèse s'est penchée sur l'exploration des domaines grisés. 165

B.1 Première possibilité d'utilisation des relation spatiales réifiées d'OBOE. L'instance de la relation spatiale est entourée de rouge. 171
B.2 Deuxième possibilité d'utilisation des relation spatiales réifiées d'OBOE. L'instance de la relation spatiale est entourée de rouge. Les propriétés incohérentes sont également en rouge. 172
B.3 Troisième possibilité d'utilisation des relation spatiales réifiées d'OBOE. Les instances de la relation spatiale sont entourées de rouge. 173
B.4 Troisième possibilité d'utilisation des relation spatiales réifiées d'OBOE déclinée pour les relations spatiales symétriques. L'instance de la relation spatiale est entourée de rouge. 174

Liste des tableaux

2.1 Les bandes spectrales des images Landsat 7 (ETMP). 31
2.2 Les bandes spectrales des images Spot 5 (HRG). 32

3.1 Les axiomes terminologiques (C_1 et C_2 sont des concepts). 70
3.2 Les constructeurs de la logique $\mathcal{FL}$ (C_i est un concept, R un rôle). 73
3.3 La négation atomique $\mathcal{AL}$ (C est un concept atomique). 73
3.4 De $\mathcal{AL}$ à $\mathcal{S}$ (C_i est un concept, R un rôle). 73
3.5 Constructeurs courants (C_i sont des concepts, R_i des rôles). 74
3.6 Quelques logiques descriptives et leur classe de complexité. 75

7.1 Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et l'approche par seuillage de pixels (comparaison pixel à pixel des images en FIGURE 7.6). 133
7.2 Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et l'approche par seuillage de pixels (comparaison pixel à pixel des images en FIGURE 7.7). 133
7.3 Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et l'approche par seuillage de pixels (comparaison pixel à pixel des images en FIGURE 7.8). 134
7.4 Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et une interprétation réalisée manuellement par Mme Frédérique SEYLER. 137
7.5 Matrice de confusion et statistiques dérivées comparant l'approche par raisonnement sur les objets et une interprétation réalisée manuellement par M. Jean-François FAURE. 137
7.6 Matrice de confusion et statistiques dérivées comparant les interprétations réalisées manuellement par Mme Frédérique SEYLER et M. Jean-François FAURE. 138
7.7 Comparaison de temps d'exécution mesurés avec et sans parallélisation du calcul des relations spatiales. Les 1161 objets sont issus d'un extrait d'image 100×100 pixels. Les 5751 objets sont issus d'un extrait d'image 200×200 pixels. Cette étape fait partie de la description de l'image (instanciation des faits de l'image). Le raisonnement, lui, laissé au raisonneur, n'est pas concerné par la parallélisation. 140
7.8 Comparaison du temps d'exécution du raisonnement sur différents extraits d'image. La colonne (A) correspond à une classification dans une base de connaissances qui contient l'ensemble des pixels d'un extrait. Dans la colonne (B) chaque pixel est classifié dans une base de connaissances distincte. 143
7.9 Comparaison des temps d'exécution de différentes déclinaisons de l'approche. . . . 145

A.1 Quelques indices spectraux implémentés dans *Orfeo Toolbox*. 169

Bibliographie

Abdellaoui, A. and Rougab, A. (1997). Caractérisation de la réponse du bâti : application au complexe urbain de blida (algérie). *Universités francophones. Actualité scientifique)*, pages 75–83.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Commun. ACM*, 26(11) :832–843.

Andrés, S. (2011). Interprétation sémantique des images satellitaires. In *Actes de la conférence Informatique pour les Systèmes d'Information de Décision*, pages 421–422, Lille, France.

Andrés, S. (2012). Génie ontologique et raisonneurs : application aux images satellitaires. In *Actes de la conférence Informatique pour les Systèmes d'Information de Décision*, pages 575–582, Montpellier, France. Prix du meilleur article doctoral.

Andrés, S., Arvor, D., Laporte, M.-A., Durieux, L., Libourel, T., Mougenot, I., and Pierkot, C. (2012a). Ontologies contribution to link thematic and remote-sensing knowledge : preliminary discussions. In *Proceedings of the International Conference on Geographic Object Based Image Analysis*.

Andrés, S., Arvor, D., and Pierkot, C. (2012b). Towards an ontological approach for classifying remote sensing images. In *SITIS 2012*.

Andrés, S., Arvor, D., and Pierkot, C. (2012c). Une approche ontologique d'aide à la classification des objets d'une image satellitaire. In *Proceedings of the Conference on Spatial Analysis and Geomatics*.

Andrés, S., Pierkot, C., and Arvor, D. (2013). Towards a semantic interpretation of satellite images by using geographic standards : focus on spatial relations. In *Geoprocessing*.

Arvor, D., Durieux, L., Andrés, S., and Laporte, M.-A. (2013a). Advances in geographic object-based image analysis with ontologies : A review of main contributions and limitations from a remote sensing perspective. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 82(0) :125 – 137.

Arvor, D., Kosmidou, V., Libourel, T., Adamo, M., Tarantino, C., Lucas, R., O'Connor, B., Blonda, P., Pierkot, C., Fargette, M., Andrés, S., and Durieux, L. (2012). Semantic nets for object-oriented land cover mapping : a preliminary example.

Arvor, D., Saint-Geours, N., Dupuy, S., Andrés, S., and Durieux, L. (2013b). Identifying optimal classification rules for geographic object-based image analysis. In *XVI Simpósio Brasileiro de Sensoriamento Remoto*, pages p. 2290 – p. 2297, Igauçu, Brésil.

Athanasiadis, T., Mylonas, P., Avrithis, Y., and Kollias, S. (2007). Semantic image segmentation and object labelling. *IEEE Transactions on circuits and systems for video techonology*, 17(3) :298–312.

Baader, F. (1996). A formal definition for the expressive power of terminological knowledge representation languages. *Journal of logic and computation*, 6(1) :33–54.

Baader, F., Bürckert, H.-J., Heinsohn, J., Müller, J., Hollunder, B., Nebel, B., Nutt, W., and Profitlich, H.-J. (1990). Terminological knowledge representation : A proposal for a terminological logic. DFKI Technical Memo TM-90-04, Deutsches Forschungszentrum für Künstliche Intelligenz, Kaiserslautern. Updated version, taking into account the results of a discussion at the "International Worksop on Terminological Logics," Dagstuhl, May 1991.

Baader, F. and Hollunder, B. (1991). A terminological knowledge representation system with complete inference algorithms. In *Proceedings of the First International Workshop on Processing Declarative Knowledge*, volume 572 of *Lecture Notes in Computer Science*, pages 67–85, Kaiserslautern (Germany). Springer–Verlag.

Baader, F., Horrocks, I., and Sattler, U. (2007). Description logics. In van Harmelen, F., Lifschitz, V., and Porter, B., editors, *Handbook of Knowledge Representation*, pages 135–179. Elsevier.

Baader, F., Milicic, M., Lutz, C., Sattler, U., and Wolter, F. (2005). Integrating description logics and action formalisms for reasoning about web services. LTCS-Report LTCS-05-02, Chair for Automata Theory, Institute for Theoretical Computer Science, Dresden University of Technology, Germany. See http ://lat.inf.tu-dresden.de/research/reports.html.

Baader, F. and Nutt, W. (2003). Basic description logics. In Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and F. Patel-Schneider, P., editors, *The Description Logic Handbook : Theory, Implementation, and Applications*, pages 43–95. Cambridge University Press.

Baader, F. and Sattler, U. (1998). Description logics with concrete domains and aggregation. In Prade, H., editor, *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI-98)*, pages 336–340. John Wiley & Sons Ltd.

Baader, F. and Sattler, U. (2003). Description logics with aggregates and concrete domains. *Information Systems*, 28(8) :979–1004.

Bail, S., Parsia, B., and Sattler, U. (2010). Justbench : a framework for owl benchmarking. In *Proceedings of the 9th international semantic web conference on The semantic web - Volume Part I*, ISWC'10, pages 32–47, Berlin, Heidelberg. Springer-Verlag.

Baraldi, A., Puzzolo, V., Blonda, P., and Bruzzone, L. (2006). Automatic spectral rule-based preliminary mapping of calibrated landsat tm and etm+ images. *IEEE Transactions on geoscience and remote sensing*, 44(9) :2563 – 2586.

Baret, F. and Guyot, G. (1991). Potentials and limits of vegetation indices for lai and apar assessment. *Remote sensing of environment*, 35(2) :161–173.

Batsakis, S. and Petrakis, E. G. M. (2010). Representing and reasoning over spatio-temporal information in owl 2.0. In *Proceedings of the 6th Workshop on Semantic Web Applications and Perspectives (SWAP' 2010)*, Bressanone, Italy.

Bayoudh, M., Roux, E., Nock, R., Richard, G., et al. (2012). Automatic learning of structural knowledge from geographic information for updating land cover maps. *Symposium of the Latin American Society for Remote Sensing and Spatial Information Systems (SELPER)*.

Ben Mohamed, K. (2010). *Traitement de requêtes conjonctives avec négation : algorithmes et expérimentations*. These, Université Montpellier II - Sciences et Techniques du Languedoc.

Bertrand, A. (2013). Optimisation de la segmentation automatique d'images satellitaires pour la cartographie dynamique : Application au littoral guyanais. Master's thesis, Université de Caen.

Bissol, S. (2006). *Indexation symbolique d'images : une approche basée sur l'apprentissage non supervisé de régularités*. PhD thesis, Université Joseph Fourier, Grenoble 1.

Bock, J., Haase, P., Ji, Q., and Volz, R. (2008). Benchmarking OWL Reasoners. *Workshop on Advancing Reasoning on the Web : Scalability and Commonsense*.

Bonn, F. (1994). *Télédétection de l'environnement dans l'espace francophone*. Presses de l'Université du Québec.

Bonnet, P. (2005). Images couleur. Cours.

Bonnet, P. (2010). Discrétisation de l'image. Cours.

Bournez, O., Dowek, G., Gilleron, R., Grigorieff, S., Marion, J.-Y., Perdrix, S., and Tison, S. (2010). Décidabilité et Complexité. In Pierre Marquis, O. P. and Prade., H., editors, *IA Handbook*, pages 1–63. Cépaduès.

Boussaid, O., Sheeren, D., Puissant, A., and Gançarski, P. (2006). Développement d'une ontologie pour la classification d'images en télédétection. In *Actes SAGEO 2006*.

Brachman, R. J. and Levesque, H. J. (1984). The tractability of subsumption in frame-based description languages. In *Proc. of the 4th Nat. Conf. on Artificial Intelligence (AAAI-84)*, pages 34–37.

Brisson, R., Boussaid, O., Gançarski, P., Puissant, A., and Durand, N. (2007). Navigation et appariement d'objets géographiques dans une ontologie. In *Actes des 7èmes Journées Francophones "Extraction et gestion des connaissances"*, pages 391–396.

Casati, R., Smith, B., and Varzi, A. (1998). Ontological Tools for Geographic Representation. In Guarino, N., editor, *Formal Ontology in Information Systems*, pages 77–85. Amsterdam : IOS Press.

Castilla, G. (2003). *Object-oriented analysis of Remote Sensing images for land cover mapping : conceptual foundations and a segmentation method to derive a baseline partition for classification.* PhD thesis, Polythecnic University of Madrid., Madrid,Spain.

Castilla, G. and Hay, G. J. (2008). *Object-based image analysis : spatial concepts for knowledge-driven remote sensing applications*, chapter Image objects and geographic objects, pages 91–110. Springer.

Christodoulou, G., Petrakis, E. G. M., and Batsakis, S. (2012). Qualitative spatial reasoning using topological and directional information in owl. In *ICTAI*, pages 596–602. IEEE.

Clementini, E. (2009). *A Conceptual Framework for Modelling Spatial Relations*. Thèse de doctorat en informatique, INSA Lyon.

Clevers, J. (1988). The derivation of a simplified reflectance model for the estimation of leaf area index. *Remote Sensing of Environment*, 25(1) :53–69.

Cohen, J. et al. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1) :37–46.

Cousin, V. (1840). *Œuvres de Victor Cousin, Cours de l'Histoire de la Philosophie*. Œuvres de Victor Cousin. Société Belge de Librairie.

Crippen, R. E. (1990). Calculating the vegetation index faster. *Remote Sensing of Environment*, 34(1) :71–73.

Dasiopoulou, S. and Kompatsiaris, I. (2010). Trends and issues in description logics frameworks for image interpretation. In *Artificial Intelligence : Theories, Models and Applications*, pages 61–70.

Derivaux, S., Forestier, G., Wemmert, C., and Lefevre, S. (2008). Extraction de détecteurs d'objets urbains à partir d'une ontologie. In *Atelier Extraction de Connaissance à partir d'Images (ECOI), Journées Francophones Extraction et Gestion des Connaissances (EGC 2008)*, pages 71–81, Sophia Antipolis, France.

Derivaux, S., Forestier, G., Wemmert, C., and Lefèvre, S. (2010). Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation. *Pattern Recognition Letters*, 31 :2364–2374.

Deschamps, C., Moulin, F., and Warusfel, A. (1999). *Mathématiques 1re année : cours et exercices corrigés*. J'intègre. Série E. Ramis. Dunod.

Doherty, P., Lukaszewicz, W., and Szalas, A. (2000). Efficient reasoning using the local closed-world assumption. In *Artificial Intelligence : Methodology, Systems, and Applications*, pages 49–58. Springer.

Durand, N., Derivaux, S., Forestier, G., Wemmert, C., Gançarski, P., Boussaïd, O., and Puissant, A. (2007). Ontology-based object recognition for remote sensing image interpretation. *19th IEEE International Conference on Tools with Artificial Intelligence*, 19 :472 – 479.

Egenhofer, M. (1989). A formal definition of binary topological relationships. In Litwin, W. and Schek, H.-J., editors, *Foundations of Data Organization and Algorithms*, volume 367 of *Lecture Notes in Computer Science*, pages 457–472. Springer Berlin Heidelberg.

Espinasse, B. (2009). Introduction aux logiques de description. Université d'Aix-Marseille.

Falomir, Z., Jiménez-Ruiz, E., Escrig, M. T., and Museros, L. (2011). Describing images using qualitative models and description logics. *Spatial Cognition & Computation*, 11(1) :45–74.

Féty, L. and Terré, M. (1991). Amelioration d'images par egalisation de capteurs. In *13e Colloque sur le traitement du signal et des images, FRA, 1991*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.

Forestier, G., Puissant, A., Wemmert, C., and Gançarski, P. (2012). Knowledge-based region labeling for remote sensing image interpretation. *Computers, Environment and Urban Systems*, 36(5) :470 – 480.

Forestier, G., Wemmert, C., and Puissant, A. (2013). Coastal image interpretation using background knowledge and semantics. *Computers & Geosciences*, 54(0) :88 – 96.

Fournier-Viger, P. (2005). Une introduction aux logiques de description. D'après *Un modèle de représentation des connaissances à trois niveaux de sémantique pour les systèmes tutoriels intelligents*, Université de Sherbrooke, Sherbrooke, Canada.

Gao, B.-C. (1996). Ndwi – a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3) :257–266.

Gardiner, T., Tsarkov, D., and Horrocks, I. (2006). Framework for an automated comparison of description logic reasoners. In Cruz, I. F., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., and Aroyo, L., editors, *International Semantic Web Conference*, volume 4273 of *Lecture Notes in Computer Science*, pages 654–667. Springer.

Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., and Sattler, U. (2008). Owl 2 : The next step for owl. *Web Semantics : Science, Services and Agents on the World Wide Web*, 6(4) :309 – 322. <ce :title>Semantic Web Challenge 2006/2007</ce :title>.

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2) :199–220.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6) :610–621.

Harold, E. and Means, W. (2002). *XML in a nutshell : manuel de référence*. O'Reilly Editions.

Hassan, S., Hétroy, F., and Palombi, O. (2009). Segmentation de maillages guidée par une ontologie. In *22èmes journées de l'association francophone d'informatique graphique*, Arles.

Herring, J. (2011). Opengis implementation standard for geographic information-simple feature access-part 1 : Common architecture.

Hichem Bannour, C. H. (2011). Towards ontologies for image interpretation and annotation. In *Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on*, Madrid, Espagne.

Hogenboom, F., Frasincar, F., and Kaymak, U. (2010). A review of approaches for representing rcc8 in owl. In Shin, S. Y., Ossowski, S., Schumacher, M., Palakal, M. J., and Hung, C.-C., editors, *SAC*, pages 1444–1445. ACM.

Horrocks, I., Patel-Schneider, P. F., McGuinness, D. L., and Welty, C. A. (2007). OWL : a Description Logic Based Ontology Language for the Semantic Web. In Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Patel-Schneider, P. F., editors, *The Description Logic Handbook : Theory, Implementation, and Applications (2nd Edition)*, chapter 14. Cambridge University Press.

Hudelot, C., Atif, J., and Bloch, I. (2008). Fuzzy spatial relation ontology for image interpretation.

Hudelot, C. and Thonnat, M. (2003). A cognitive vision platform for automatic recognition of natural complex objects. In *Tools with Artificial Intelligence, 2003. Proceedings. 15th IEEE International Conference on*, pages 398–405. IEEE.

Huete, A., Justice, C., and Liu, H. (1994). Development of vegetation and soil indices for modis-eos. *Remote Sensing of Environment*, 49(3) :224–234.

Huete, A. R. (1988). A soil-adjusted vegetation index (savi). *Remote sensing of environment*, 25(3) :295–309.

Hustadt, U. (1994). Do we need the closed-world assumption in knowledge representation ?

Hyvönen, E., Styrman, A., and Saarela, S. (2002). Ontology-based image retrieval. pages 15–27.

Iakovidis, D., Schober, D., Boeker, M., and Schulz, S. (2009). An ontology of image representations for medical image mining. In *Information Technology and Applications in Biomedicine, 2009. ITAB 2009. 9th International Conference on*, pages 1 –4.

Katz, Y. and Grau, B. C. (2005). Representing qualitative spatial information in owl-dl. In *Proceedings of the OWL : Experiences and Directions Workshop. Galway*.

Kaufman, Y. J. and Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2) :261–270.

Kazakov, Y. (2008). Sriq and sroiq are harder than shoiq. In *Description Logics*.

Klein, É. (2007). *Le facteur temps ne sonne jamais deux fois*. Nouvelle bibliothèque scientifique. Flammarion.

Kolas, D. and Self, T. (2007). Spatially-augmented knowledgebase. In Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., and Cudré-Mauroux, P., editors, *The Semantic Web*, volume 4825 of *Lecture Notes in Computer Science*, pages 792–801. Springer Berlin Heidelberg.

Koletsis, P. (2009). *Semantic Image annotation using Ontologies and Image Content Analysis*. PhD thesis, Technical University of Crete.

Koubarakis, M., Karpathiotakis, M., Kyzirakos, K., Nikolaou, C., and Sioutis, M. (2012). Data models and query languages for linked geospatial data. In Eiter, T. and Krennwallner, T., editors, *Reasoning Web. Semantic Technologies for Advanced Query Answering*, volume 7487 of *Lecture Notes in Computer Science*, pages 290–328. Springer Berlin Heidelberg.

Lacaux, J., Tourre, Y., Vignolles, C., Ndione, J., and Lafaye, M. (2007). Classification of ponds from high-spatial resolution remote sensing : Application to rift valley fever epidemics in senegal. *Remote Sensing of Environment*, 106(1) :66–74.

Landgrebe, D. (1997). Multispectral data analysis : A signal theory perspective.

Laporte, M.-A. (2011). *Définition de standards de données relatives aux traits fonctionnels des végétaux pour l'étude de la biodiversité*. PhD thesis, SIBAGHE.

Lehmann, G. (2007). Label object representation and manipulation with itk.

Levin, M. K., Ruttenberg, A., Masci, A. M., and Cowell, L. G. (2011). owl_cpp, a c++ library for working with owl ontologies. In Bodenreider, O., Martone, M. E., and Ruttenberg, A., editors, *ICBO*, volume 833 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Lin, Y. (2011). *Méthodologie et composants pour la mise en oeuvre de workflows scientifiques*. PhD thesis, Information, Structures, Systèmes, Université Montpellier 2.

Liu, Y., Guo, Q., and Kelly, M. (2008). A framework of region-based spatial relations for non-overlapping features and its applications in object based image analysis. *ISPRS Journal of Photogrammetry & Remote Sensing*, 63 :461–475.

Maillot, N. and Thonnat, M. (2008). Ontology based complex object recognition. *Image Vision Comput.*, 26(1) :102–113.

Maillot, N., Thonnat, M., and Hudelot, C. (2004). Ontology based object learning and recognition : application to image retrieval. In *Tools with Artificial Intelligence*, pages 620–625.

Manzanera, A. (2011). Les images numériques. Cours.

Mariño Drews, O. (1993). *Raisonnement classificatoire dans une représentation à objets multi-points de vue*. PhD thesis, Université Joseph-Fourier-Grenoble I.

McFeeters, S. (1996). The use of the normalized difference water index (ndwi) in the delineation of open water features. *International journal of remote sensing*, 17(7) :1425–1432.

Mezaris, V., Kompatsiaris, I., and Strintzis, M. G. (2003). An ontology approach to object-based image retrieval. In *In Proc. IEEE Int. Conf. on Image Processing (ICIP03)*, pages 511–514.

Minsky, M. (1974). A framework for representing knowledge.

Nardi, D. and Brachman, R. J. (2003). An introduction to description logics. In Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and F. Patel-Schneider, P., editors, *The Description Logic Handbook : Theory, Implementation, and Applications*, pages 5–44. Cambridge University Press.

Neumann, B. and Möller, R. (2008). On scene interpretation with description logics. *Image and Vision Computing*, 26 :82–101.

OWL-a (2012). Owl 2 web ontology language document overview.

OWL-d (2012). Owl 2 web ontology language direct semantics.

OWL-h (2012). Owl 2 web ontology language primer.

Pan, J. Z. (2004). *Description Logics : Reasoning Support for the Semantic Web*. PhD thesis, School of Computer Science, The University of Manchester, Oxford Rd, Manchester M13 9PL, UK.

Pearson, R. L. and Miller, L. D. (1972). Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie. In *Remote Sensing of Environment, VIII*, volume 1, page 1355.

Pierkot, C., Andrés, S., Seyler, F., and Faure, J.-F. (2013). Formalizing spatio-temporal knowledge in remote sensing applications to improve image interpretation. *Journal of Spatial Information Science*.

Platon (1822). *Œuvres de Platon, traduites par Victor Cousin, Phédon, ou de l'âme*. Œuvres de Platon.

Pouget, F. (2005). *Notions de base de télédétection*. Université de La Rochelle, France.

Pouget, M., Madeira, J., Le Floch, E., and Kamal, S. (1990). Caracteristiques spectrales des surfaces sableuses de la region cotiere nord-ouest de l'egypte : application aux donnees satellitaires spot..

Puissant, A., Durand, N., Sheeren, D., Weber, C., and Gançarski, P. (2007). Urban Ontology for Semantic Interpretation of Multi-Source Images. In *2nd Workshop Ontologies for urban development : conceptual models for practitioners (Urban Ontologies 2007)*, Castello del Valentino, Turin, Italy.

Puissant, A., Sheeren, D., Weber, C., Wemmert, C., and Gançarski, P. (2006). Amélioration des connaissances sur l'environnement urbain : intérêt de l'intégration de règles dans les procédures de classification. In *Interactions Nature-Société, analyse et modèles*.

Qi, J., Chehbouni, A., Huete, A., Kerr, Y., and Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote sensing of environment*, 48(2) :119–126.

Quillian, M. R. (1967). Word concepts : a theory and simulation of some basic semantic capabilities. *Behavioral Science*, 12(5) :410–430.

Randell, D. A., Cui, Z., and Cohn, A. G. (1992). A Spatial Logic based on Regions and Connection. In *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*.

RDF-a (2004). Rdf primer.

RDF-e (2004). Rdf vocabulary description language 1.0 : Rdf schema.

Reix, R. (2004). *Systèmes d'information et management des organisations*. Vuibert gestion. Vuibert.

Richardson, A. J. and Weigand, C. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43(12).

Rouse, J., Haas, R., Schell, J., and Deering, D. (1973). Monitoring vegetation systems in the great plains with erts. pages 309–317. cited By (since 1996)1.

Ruan, N., Huang, N., and Hong, W. (2006). Semantic-Based Image Retrieval in Remote Sensing Archive : An Ontology Approach. In *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, pages 2903–2906.

Rudolph, S. (2011). Foundations of description logics. In *Reasoning Web*, pages 76–136.

Saint-Dizier, P. (2006). Réseaux sémantiques. In Godard, D., Roussarie, L., and Corblin, F., editors, *Sémanticlopédie : dictionnaire de sémantique*. GDR Sémantique & Modélisation, CNRS, http ://www.semantique-gdr.net/dico/.

Sanz, S., Régagnon, E., and Barge, O. (2008).

Schmidt-Schauss, M. and Smolka, G. (1991). Attributive concept descriptions with complements. *Artificial Intelligence*, 48 :1–26.

Schober, J.-P., Hermes, T., and Herzog, O. (2004). Content-based image retrieval by ontology-based object recognition. In *in KI-2004 Workshop on Applications of Description Logics*, pages 61–67.

Sebari, I. and He, D.-C. (2007). Les approches de segmentation d'image par coopération régions-contours. *Revue télédétection*, 7(1-2-3-4) :499–506.

Shadbolt, N. and Burton, M. (1995). Knowledge elicitation : a systematic approach. *Evaluation of human work : a practical ergonomics methodology*, pages 406–440.

Sheeren, D., Quirin, A., Puissant, A., Gancarski, P., and Weber, C. (2006). Discovering rules with genetic algorithms to classify urban remotely sensed data. In *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, pages 3919 –3922.

Simari, P., Nowrouzezahrai, D., Kalogerakis, E., and Singh, K. (2009). Multi-objective shape segmentation and labeling. In *Computer Graphics Forum : Eurographics Symposium on Geometry Processing*, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Smeulders, A. W. N., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12) :1349–1380.

Sowa, J. F. (1991). Toward the expressive power of natural language. *Principles of Semantic Networks*, pages 157–189.

SPOT-a (2010). Le format dimap. Technical report, Spot Image.

SPOT-b (2010). Niveaux de prétraitement et précision de localisation. Technical report, Spot Image.

Srikanth, M., Varner, J., Bowden, M., and Moldovan, D. (2005). Exploiting ontologies for automatic image annotation. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '05, pages 552–558, New York, NY, USA. ACM.

Studer, R., Benjamins, V. R., and Fensel, D. (1998). Knowledge engineering : principles and methods. *Data & knowledge engineering*, 25(1) :161–197.

Team, O. D. (2010). *OTB Software Guide*. Centre National d'Études Spatiales.

Tobies, S. (2001). *Complexity Results and Practical Algorithms for Logics in Knowledge Representation*. PhD thesis, LuFG Theoretical Computer Science, RWTH-Aachen, Germany.

Tsarkov, D. and Horrocks, I. (2006). Fact++ description logic reasoner : System description. In *Proc. of the Int. Joint Conf. on Automated Reasoning (IJCAR 2006)*, volume 4130 of *Lecture Notes in Artificial Intelligence*, pages 292–297. Springer.

Verbeke, G. (1952). La doctrine de l'être dans la métaphysique d'aristote. *Revue Philosophique de Louvain*, 50(27) :471–478.

Weithöner, T., Liebig, T., Luther, M., and Böhm, S. (2006). What's Wrong with OWL Benchmarks ? In *Proc. of the Second Int. Workshop on Scalable Semantic Web Knowledge Base Systems (SSWS 2006)*, pages 101–114, Athens, GA, USA.

Wu, K. and Haarslev, V. (2012). A parallel reasoner for the description logic alc. In Kazakov, Y., Lembo, D., and Wolter, F., editors, *Description Logics*, volume 846 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Xiaoguang, F. (2002). Content-based image retrieval. Based on : J. Huang. Color-Spatial Image Indexing and Applications. Ph.D thesis, Cornell Univ., 1998.

XML-a (2008). Extensible markup language (xml) 1.0 (fifth edition).

Xu, H. (2006). Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14) :3025–3033.

Zha, Y., Gao, J., and Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from tm imagery. *International Journal of Remote Sensing*, 24(3) :583–594.

Zhang, X., Xiao, P., and Feng, X. (2012). An unsupervised evaluation method for remotely sensed imagery segmentation. *IEEE Geoscience and remote sensing letters*, 9(2) :156–160.