



HAL
open science

Proposition d'une annotation sémantique floue guidée par ontologie pour l'interprétation des images de télédétection: Application à la gestion des risques naturels

Wassim Messoudi

► To cite this version:

Wassim Messoudi. Proposition d'une annotation sémantique floue guidée par ontologie pour l'interprétation des images de télédétection: Application à la gestion des risques naturels. Traitement des images [eess.IV]. Télécom Bretagne, Université de Rennes 1, 2013. Français. NNT: . tel-00833902

HAL Id: tel-00833902

<https://theses.hal.science/tel-00833902>

Submitted on 13 Jun 2013

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.

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Rennes 1

Co-tutelle avec l'Ecole Nationale des Sciences de l'Informatique,
Université de la Manouba

Ecole Doctorale – MATISSE

Proposition d'une annotation sémantique floue guidée par ontologie pour l'interprétation des images de télédétection: Application à la gestion des risques naturels

Thèse de Doctorat

Mention : Informatique

Présentée par **Wassim Messoudi**

Département : ITI, Télécom Bretagne
&

Laboratoire : RIADI, Ecole Nationale des Sciences de l'Informatique

Directeurs de thèse : M. Basel Solaiman
M. Imed Riadh Farah

Soutenue le 23 janvier 2013

Jury :

M. Ali Khenchaf, Professeur à l'ENSTA Bretagne (Président)
M. Jean-Paul Haton, Professeur Université Henri Poincaré, Nancy 1 (Rapporteur)
M. Jalel Akaichi, Maître de conférences (HDR) à l'ISG (Rapporteur)
M. Robert Laurini, Professeur émérite, Insa de Lyon (Examineur)
M. Basel Solaiman, Professeur à Télécom Bretagne (Directeur de thèse)
M. Imed Riadh Farah, Maître de conférences (HDR) à l'ISAMM (Directeur de thèse)

Résumé

L'image de télédétection constitue une source d'information importante permettant de surveiller les phénomènes naturels auxquels le monde a été confronté ces dernières années notamment l'érosion, l'inondation, les incendies, la déforestation, etc. Ces phénomènes ont des conséquences graves sur l'environnement, l'agriculture (dégradation des sols), l'activité humaine, etc. C'est dans ce cadre que se situe notre travail de thèse qui s'intéresse à l'interprétation des images de télédétection afin d'étudier la susceptibilité des scènes aux risques naturels.

Cette thèse aborde deux problématiques majeures; La première concerne la modélisation du contenu de l'image satellitale. Cela pose les questions suivantes : Comment extraire les informations utiles de l'image ? Quel modèle de représentation choisir pour modéliser l'image ? Comment générer cette modélisation ? La deuxième problématique concerne l'étude de la susceptibilité de l'image aux risques naturels en se basant sur les sciences cognitives et les connaissances issues de l'image. Donc, quel est le modèle de raisonnement à adopter pour générer une interprétation adéquate ?

Dans ce travail de thèse, nous proposons une ontologie spatio-temporelle, floue, lourde et évolutive représentant le contenu sémantique des images de télédétection. Pour ce faire, nous avons suivi un processus de développement basé sur la réutilisation des ressources ontologiques existantes, l'apprentissage à partir des images satellitales (de référence), et la formalisation des connaissances a priori de domaine (risques naturels, connaissance sur les scènes capteurs d'acquisition, etc.).

Nous proposons également une méthodologie permettant de générer des annotations modélisant le contenu de l'image satellitale en se guidant par l'ontologie de domaine, et permettant, à partir d'un raisonnement inférentiel sur ces connaissances, de déduire la susceptibilité d'une scène satellitale aux risques naturels. La méthodologie proposée comporte quatre modules : (1) Annotation de l'image de télédétection, (2) Fusion d'annotations, (3) Recherche de cas, et (4) Raisonnement et interprétation.

Dans un cadre méthodologique, nous avons appliqué la méthodologie pour traiter le problème d'érosion afin de prévenir les zones à risque en Tunisie et plus particulièrement à la région du Kef. Pour ce faire, nous avons conçu une base d'images multi-dates et multi-capteurs représentant des classes d'ontologie (forêt, lac, champ de blé, zone urbaine, ...). De même, nous avons construit une base de cas contenant des annotations qui représentent des images de télédétection de la région du Kef. Dans un processus « off-line », nous avons simulé le processus d'interprétation sur des requêtes pour voir si elles présentent des risques d'érosion.

Mots clés : *Imagerie satellitale, ontologie, enrichissement d'ontologie, annotation sémantique, raisonnement à base cas.*

Abstract

The remotely sensed images are considered as an important source of information used in several domains such as environmental monitoring, disaster management and military intelligence. However, the appropriate information has to be extracted and represented to make efficient decision processes.

In this thesis, we proposed a spatiotemporal, fuzzy and heavy ontology for modeling remote sensing images. The ontology is based on (1) the re-use of an existing ontology of domain which is considered as a kernel-ontology, and (2) the enrichment of this one in order to cover the domain of application and to take into account the specificities of remotely sensed images. The process of enrichment establish at various levels: conceptual, relational and axiomatic, basing on external resources like taxonomies, thesaurus, and knowledge extracted from experts and image archive. The proposed enrichment allows maximizing the capacity of the ontology on modeling and reasoning through the axioms. We propose to enrich the ontology with fuzzy logic in order to modeling the uncertainty in remotely sensed images.

Furthermore, we propose a methodology for modeling and retrieving satellite images basing on their spatial knowledge. The main idea of our approach is that the use of spatial knowledge, reasoning and inference technique, can contribute to deduce the susceptibility of the scene at natural risks (erosion, flooding, fires, etc.).

Our methodology takes in input a set of multi-sensor images representing a scene. It contains four modules: (1) Modeling of the scene, (2) fusion of image annotations, (3) similar case retrieval, and (4) reasoning and interpretation. The first module generates annotations which represents the semantic content of the satellite scene. The second module allows merging image annotations to have faithful information to the reality. The third module attempts to find similar cases to those of the annotated query to take advantage of cases, situations and past problems. The fourth module allows deducting the susceptibility of the image in a given natural phenomenon (Erosion, flooding, fires), basing on inferences on knowledge of domain (Ontology, knowledge of expert, knowledge of natural phenomena, etc.) and the result of the previous phases.

Finally, we applied the methodology to deal with the problem of erosion to warn high-risk areas of the kef region, situated in northwest of Tunisia. Indeed, we developed a base of images representing the classes of the ontology (forest, lake, urban zone, etc.). Also, we developed a base of cases representing remote sensing images of the region of the Kef. In an "offline" process, we simulated the process of interpretation on requests to see if they present risks of erosion.

Keywords: *Remote sensing images, ontology, ontology enrichment, semantic annotation, case based-reasoning.*

Je dédie mes travaux de thèse

à ma chère femme Thouraya et ma chère fille Ayouta ;

à mes chers parents Sassi et Kalthoum ;

à mes chers frères Rached et housseem ;

à toute ma famille et ma belle-famille ;

et à tous mes amis ;

Remerciements

Je tiens tout d'abord à exprimer ma reconnaissance et ma gratitude à mes directeurs de thèse : Monsieur Imed Riadh FARAH, maître de conférences (HDR) à l'Université de la Mannouba, et Monsieur Basel SOLAIMAN, professeur et directeur du département Image et Traitement de l'Information, pour la qualité de leurs encadrements et leurs orientations tout au long de l'élaboration de cette thèse.

Je voudrais de même remercier Monsieur Mohamed Farah, Maître assistant à l'Institut Supérieur des Arts Multimédias de la Mannouba pour sa disponibilité, ses précieux conseils et son encouragement durant ce travail de recherche.

Mes remerciements s'adressent également à Monsieur Ali KHENCHAF, Professeur à l'ENSTA Bretagne, pour l'honneur qu'il m'a accordé en présidant mon jury de thèse.

Je tiens à exprimer ma profonde reconnaissance à Monsieur Jean Paul HATON, Professeur à l'Université Henri Poincaré Nancy 1, et Monsieur Jalel Akaichi, Maître de conférences (HDR) à l'Institut Supérieur de Gestion, d'avoir accepté la fastidieuse tâche de rapporter ce travail. Je les remercie pour les remarques constructives qu'ils m'ont permis de perfectionner le présent manuscrit.

Je remercie également Monsieur Robert Laurini Professeur émérite à l'Institut National des Sciences Appliquées de Lyon, qui m'a fait plaisir d'examiner ce travail.

Mes remerciements vont également à tous les membres du laboratoire RIADI et du département ITI et à tous ceux qui m'ont aidé et encouragé tout au long de cette thèse, et plus particulièrement à Monsieur Karim Saheb Ettabaa pour ces précieux conseils.

Enfin, J'aimerais exprimer toute ma gratitude à ma femme pour sa patience et son soutien tout le long de ce travail. Merci également à mes parents pour leurs encouragements dès le début.

Table des matières

Introduction générale	1
-----------------------------	---

Partie A : Etat de l'art

Chapitre 1 : Annotation sémantique d'images de télédétection

1. L'image de télédétection	8
1.1. Définition d'une image de télédétection	8
1.2. Principe d'acquisition : la télédétection	9
1.3. Les capteurs d'acquisition	9
1.4. Types d'images de télédétection	10
1.4.1. L'image optique	10
1.4.2. L'image Radar	11
1.5. Prétraitement des images de télédétection	11
1.6. Spécificités de l'image de télédétection	11
1.6.1. L'image de télédétection, information liée au temps	11
1.6.2. Imperfections entachant l'image de télédétection	13
1.7. Interprétations des images satellitaires	13
2. Annotation sémantique d'images guidée par ontologie	14
2.1. L'annotation sémantique	14
2.2. Systèmes d'annotation existants	15
2.3. Synthèse et discussion	19
2.3.1. Rôle de l'ontologie dans l'annotation sémantique d'images	19
2.3.2. Composants d'un système sémantique d'annotation d'images par ontologie	19

Chapitre 2 : Représentation des connaissances de l'image de télédétection par ontologie

1. La notion d'ontologie	20
1.1. Définitions	20
1.2. Constituants d'ontologie	22
1.2.1. Les concepts	22
1.2.2. Les relations	23
1.2.3. Les axiomes	23
1.2.4. Les instances	24
1.3. Classification des ontologies	24
1.3.1. Classification selon l'objet de conceptualisation	24
1.3.2. Classification selon l'engagement sémantique	25
1.3.3. Classification selon le niveau de détail	26
1.4. Méthodologies de construction des ontologies	26

1.4.1. La méthode "Enterprise Ontology	26
1.4.2. La méthodologie On-To-Knowledge(OTK)	27
1.4.3. La méthodologie METHONTOLOGY	27
1.4.4. La méthodologie STANFORD	27
1.4.5. La méthodologie Kactus	25
1.4.6. La méthodologie TOVE	25
1.4.7. La méthodologie SENSUS	29
1.5. Intégration et alignement d'ontologie	29
1.6. Langages de représentation des ontologies	30
1.6.1. Les réseaux sémantiques	30
1.6.2. Les frames	30
1.6.3. Les logiques de description	31
1.6.4. Les graphes conceptuels	31
1.6.5. Les langages du Web sémantique	32
2. Ontologies existantes modélisant l'image de télédétection	32
2.1. Etat de l'art	33
2.1.1. L'ontologie INSEE	33
2.1.2. L'ontologie Towntology	34
2.1.3. L'ontologie de Fodomust	35
2.1.4. L'ontologie Ontogeo-V2-47	36
2.1.5. L'ontologie TOP10nl	37
2.1.6. L'ontologie Theme	38
2.1.7. L'ontologie Hydrology	39
2.1.8. L'ontologie de DAFOE	40
2.1.9. L'ontologie FusionTopoCarto2	41
2.1.10. L'ontologie FTT-01	42
2.2. Synthèse	42
2.3. Discussion	45

Partie B : Contributions à l'interprétation d'images de télédétection par ontologie

Chapitre 3 : Ontologie spectro-spatio-temporelle, floue, lourde et évolutive

1. Proposition d'une ontologie spatio-temporelle, floue, lourde et évolutive	47
2. Processus de développement de l'ontologie proposée	48
2.1. Choix de l'ontologie de noyau	49
2.2. Phase d'évolution de l'ontologie : vers une représentation spectro-spatio-temporelle et floue	51
2.3. Phase d'enrichissement de l'ontologie de noyau	52
2.3.1. Enrichissement conceptuel par réutilisation de ressources ontologiques	52
2.3.2. Enrichissement intentionnel de concepts à partir des images...	61
2.3.3. Enrichissement relationnel	61
2.3.4. Enrichissement axiomatique	64
2.4. Formalisme de l'ontologie	65

3. Présentation de l'ontologie résultante	66
3.1. Choix de ressources ontologiques candidates	66
3.2. Enrichissement conceptuel de l'ontologie de noyau	67
3.2.1. Résultats d'alignement et d'enrichissement pour l'ontologie source FTT.owl	67
3.2.2. Résultats d'alignement et d'enrichissement pour l'ontologie source hydrology.owl	71
3.2.3. Résultats finaux	75
3.3. Enrichissement intentionnel de l'ontologie de noyau	76
3.3.1. Elaboration d'une banque d'imagette	76
3.3.2. Illustration du processus d'enrichissement intentionnel	77
3.4. Enrichissement axiomatique	80

Chapitre 4 : Méthodologie d'annotation sémantique et l'interprétation de l'image de télédétection

1. Proposition d'une méthodologie d'interprétation sémantique d'images de télédétection	82
2. Annotation floue de l'image de télédétection guidée par ontologie.....	83
2.1. Processus de génération de l'annotation	83
2.2. Processus de matching flou	84
2.3. Méta-information proposée de l'image	85
3. Fusion floue des annotations d'images de télédétection	88
3.1. Processus de fusion proposé	88
4. Interprétation de l'image de télédétection à base d'ontologie et du RBC	90
4.1. Présentation du système d'interprétation proposé	90
4.2. Structure des cas	92
4.3. Remémoration guidée par ontologie	93
4.4. Similarité entre les objets	93
4.4.1. Similarité entre les relations spatiales	95
4.4.2. Similarité entre les contraintes	96
4.5. La phase d'adaptation	97
4.5.1. Le problème d'adaptation	97
4.5.2. Technique d'adaptation adoptée	98

Partie C : Application thématique

Chapitre 5 : Etude de la susceptibilité de scènes aux risques naturels, application à l'érosion

1. Les risques naturels	100
2. Application de la méthodologie	105
2.1. Le mode <i>offline</i>	105
2.2. Le mode <i>online</i>	109

Conclusions et perspectives	113
--	-----

Annexes

Annexe 1 : Les techniques d'alignement	117
Annexe 2 : Les sous ensembles flous	124
Annexe 3 : Convention de cotutelle	130

Publications	135
---------------------------	-----

Bibliographie	137
----------------------------	-----

Liste des figures

Figure I.1: Méthodologie de représentation et d'interprétation d'images satellitales	4
Figure 1.1 : Mesures de réflectance à différentes dates pour une parcelle de blé d'hiver	12
Figure 1.2 : Comportements spectraux de quelques matériaux suivant les capteurs SOPT et LandSat	12
Figure 1.3 : Ontologie d'objets	15
Figure 1.4 : Architecture de la plate-forme d'interprétation d'image	17
Figure 1.5 : Ontologie regroupant les concepts visuels et textuels	18
Figure 1.6 : Extrait de l'ontologie de Durant et al., (2007)	18
Figure 2.1 : Les différentes représentations de connaissance à partir de leur degré d'engagement sémantique	25
Figure 2.2 : Extrait de l'ontologie INSEE	33
Figure 2.3 : Extrait de l'ontologie de projet towntology	34
Figure 2.4 : Extrait de l'ontologie de Fodomust	35
Figure 2.5 : Extrait de l'ontologie d'objets géographiques OntoGeo-V2-47	36
Figure 2.6 : Extrait de l'ontologie d'objets géographiques Top10nl	37
Figure 2.7 : Types de propriétés de données du concept RoadSegment	38
Figure 2.8 : Extrait de l'ontologie Theme	39
Figure 2.9 : Extrait de l'ontologie Hydrology	40
Figure 2.10 : Extrait de l'ontologie d'objets géographiques DAFOE	41
Figure 2.11 : Extrait de l'ontologie d'objets géographiques FusionTopoCarto2	41
Figure 2.12 : Extrait de l'ontologie d'objets géographiques FTT	42
Figure 3.1 : Processus de développement d'ontologie	48
Figure 3.2 : Extrait de l'ontologie de Durand et al. (2007)	49
Figure 3.3 : Fonction d'appartenance gaussienne d'un attribut	50
Figure 3.4 : Exemples de fonctions d'appartenance gaussiennes décrivant les attributs d'un champ de blé	51
Figure 3.5 : Processus d'enrichissement conceptuel	53
Figure 3.6 : Alignement, filtrage, et placement entre O_{cible} et O_{source}	54
Figure 3.7 : Processus d'enrichissement intensionnel des concepts à partir d'images de télédétection...	62
Figure 3.8 : Exemples de «paysage flou» représentant les relations de distances et de direction	63

Figure 3.9 : Extrait de l'ontologie de relations spatiales floues (Hudelot et al., 2008)	64
Figure 3.10 : Extrait de l'ontologie résultante après l'ajout des descendants du concept Forest	71
Figure 3.11 : Extrait de l'ontologie résultante après l'ajout des descendants du concept waterfalls	71
Figure 3.12 : Quelques métriques sur l'ontologie de départ	75
Figure 3.13 : Extrait de l'ontologie résultante après Ajout du concept waterfalls et ses fils	76
Figure 3.14 : Quelques métriques sur l'ontologie après son enrichissement	76
Figure 3.15 : Extrait de la banque d'images	78
Figure 3.16 : Processus d'élaboration des fonctions d'appartenances	79
Figure 4.1: Méthodologie proposée pour l'interprétation sémantique d'images de télédétection	83
Figure 4.2 : Processus de génération d'annotation d'image de télédétection	84
Figure 4.3 : Illustration du processus de matching flou	86
Figure 4.4 : Méta-information proposée de l'image de télédétection	86
Figure 4.5 : Représentation XML d'une méta-information	87
Figure 4.6 : Processus de fusion d'annotations	89
Figure 4.7 : Architecture du système d'interprétation d'images proposé	91
Figure 4.8 : Structure des cas cible et source	93
Figure 4.9. Importance des relations spatiales dans mesure de relations spatiales	95
Figure 4.10 : Problème d'adaptation représenté par le carré d'analogie	97
Figure 5.1 : (a) Image de télédétection issues de capteur Spot en mois de mars, (b) régions extraits de l'image	106
Figure 5.2 : Extrait de l'annotation de l'image	108
Figure 5.3 : Image requête	110
Figure 5.4 : Cas représentant l'image requête	110
Figure 5.5 : Cas cible après mise à jour	111
Figure A2.1 : Ensemble flou avec une hauteur qui tend vers	125
Figure A2.2 : Illustration de la hauteur, du noyau et de support d'un ensemble flou	126
Figure A2.3 : Nombre flou trapézoïdal	126
Figure A2.4 : Nombre flou triangulaire	127
FigureA2.5 : Nombre flou avec des cotés paraboliques	127
Figure A2.6 : Distance de Hamming relative	128

Liste des tableaux

Tableau 1.1 : Caractéristiques de quelques capteurs et satellites d'observation de la terre	10
Tableau 2.1 : Synthèse sur les ontologies géographiques	44
Tableau 4.1 : Structure proposée d'un concept de l'ontologie	52
Tableau 4.2 : Relation spatiales topologiques RCC-8	63
Tableau 4.3 : Liste de mappings d'équivalence pour l'ontologie source FTT.owl	68
Tableau 4.4 : Liste de mappings de subsomption pour l'ontologie source FTT.owl	69
Tableau 4.5 : Résultats d'alignement pour quatre itérations sur l'ontologie FTT.owl	70
Tableau 4.6 : Liste de mappings d'équivalence pour l'ontologie source Hydrology.owl	72
Tableau 4.7 : Liste de mappings de subsomption pour l'ontologie source Hydrology.owl	73
Tableau 4.8 : Résultats d'alignement pour quatre itérations sur l'ontologie Hydrology.owl	74
Tableau 4.9 : Structure intensionnelle du concept « lac »	80
Tableau 5.1 : Primitives floues d'une région	106
Tableau 5.2 : Similarités floues entre les régions et les concepts ontologiques	107
Tableau 5.3 : Récapitulatif des similarités entre le cas cible et la base de cas	111

Introduction générale

1. Contexte général

Avec l'évolution massive des systèmes d'acquisition d'images satellitaires tels que les radars et les satellites optiques, un volume énorme d'images a été généré formant des grandes bases d'images de différentes natures, résolutions, types, dimensions, et représentant des différents types de scènes à savoir les scènes urbaines, agricoles, forestières, et industrielles.

L'image de télédétection constitue une source d'information importante permettant de surveiller les phénomènes naturels auxquels le monde est continuellement confronté notamment l'érosion, l'inondation, les incendies et la déforestation. En effet, ces phénomènes ont des conséquences graves sur l'environnement, l'agriculture (dégradation des sols), et l'activité humaine.

De ce fait, plusieurs travaux de recherches sont menés pour extraire l'information utile de l'image de télédétection, d'autres pour la représentation de son contenu, ainsi que son exploitation à l'étude des phénomènes naturels.

Notre travail de thèse se situe dans cette thématique. Etant donné une image décrivant une scène satellitale, nous cherchons à étudier sa susceptibilité aux risques naturels tout en tenant compte des connaissances de domaines, l'organisation spatiale et les contraintes temporelles des objets des images satellitaires.

2. Problématiques

L'interprétation des images de télédétection est une tâche indispensable pour surveiller les modifications subies par l'environnement d'une région géographique donnée, d'établir des études diachroniques et un suivi des phénomènes naturels.

Une image satellitale, représentant une scène, est très riche en information qui dépend essentiellement du type de capteur utilisé dans son acquisition (radar ou optique), et le facteur temporel qui reflète l'évolution de la scène dans temps.

Cette thèse aborde deux problématiques majeures; la première concerne la modélisation du contenu de l'image satellitale. Cela pose les questions suivantes : Comment extraire les informations utiles de l'image ? Quel modèle de représentation choisir pour modéliser l'image ? Comment générer cette modélisation en tenant compte des imperfections entachant les images ?

La deuxième problématique concerne l'étude de la susceptibilité de l'image aux risques naturels en se basant sur les sciences cognitives et les connaissances issues de l'image. Donc, quel est le modèle de raisonnement à adopter pour générer une interprétation adéquate ?

3. Principales contributions

3.1. Ontologie spectro-spatio-temporelle, floue, lourde et évolutive

Afin de modéliser le contenu sémantique des images de télédétection tout en tenant compte de ses caractéristiques spectrales et temporelles, nous avons proposé une ontologie ayant les spécificités suivantes :

Spatio-spectro-temporelle : *Pour modéliser la spatialité, la temporalité et surtout la caractéristique spectrale des images de télédétection. Cette caractéristique est très importante pour une interprétation sémantique d'une image de télédétection.*

Floue : *Afin de traiter les imperfections entachant l'image de télédétection. La définition floue concerne la définition des fonctions d'appartenance représentant les valeurs a priori des concepts de l'ontologie.*

Lourde : *Afin de doter l'ontologie des mécanismes de raisonnement inférentiel sur un domaine d'application. Cela permet de contribuer à créer des systèmes d'aides à la décision stratégique, à savoir la détection zones à risques d'érosion d'une région.*

Evolutive : *Pour qu'elle soit adaptable et applicable dans plusieurs domaines. Cela permet de réutiliser et exploiter les ressources ontologiques existantes.*

Nous avons proposé un processus de développement basé sur la réutilisation des ressources ontologiques existantes, l'exploitation des informations issues des images de télédétection, et la formalisation des connaissances *a priori* de domaine (risques naturels, connaissance sur les scènes, et le capteur d'acquisition). Ce processus est composé par ces étapes :

- 1) Choix de l'ontologie de noyau
- 2) Evolution structurelle de l'ontologie de noyau
 - a. Modélisation floue.
 - b. Ajout de la dimension temporelle et spectrale
- 3) Enrichissement conceptuel, intensionnel et relationnel de l'ontologie de noyau
 - a. Découverte des ressources ontologiques.
 - b. Evaluation : détermination des parties à réutiliser de chaque ressource en utilisant des techniques d'alignement d'ontologies.
 - c. Intégration et placement de nouveaux concepts.
- 4) Enrichissement axiomatique

Nous avons validé ce processus avec une ontologie de noyau tirée du travail de (Durand et al., 2007) dans le cadre du projet Fodumust, des ressources ontologiques dans le domaine géographique, et des connaissances géographique sur le domaine de risque naturel d'érosion. Nous avons obtenu des résultats satisfaisants pour un travail initial avec l'ontologie.

3.2. Méthodologie d'interprétation sémantique d'images de télédétection

Comme cadre méthodologique de travail, nous avons considéré cette ontologie comme support de connaissances pour (1) guider l'annotation sémantique des images de télédétection, et (2) guider le raisonnement inférentiel renforcé par le raisonnement à base de cas afin d'étudier la susceptibilité d'une scène aux risques naturels.

Nous avons proposé alors une méthodologie composée de quatre modules : (1) Annotation sémantique floue, (2) Fusion floue des annotations, (3) Recherche d'annotations, et (4) Raisonnement et interprétation (figure I.1).

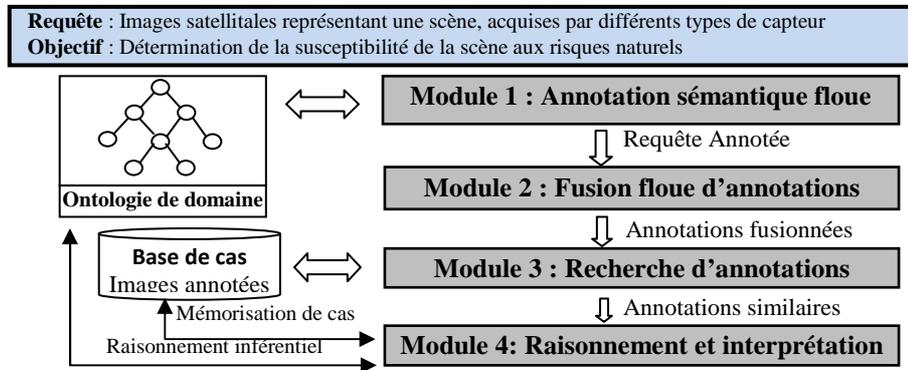


Figure I.1: Méthodologie de représentation et d'interprétation d'images satellitaires

3.2.1. Annotation sémantique floue

Ce module permet d'annoter l'image satellitale par des méta-informations décrivant son contenu en se basant sur les connaissances issues de l'image elle-même, et en se guidant par l'ontologie de domaine proposée.

Nous avons alors proposé un processus d'annotation permettant de générer une méta-information modélisant le contenu sémantique d'une image. Ce processus est basé sur la similarité floue entre les régions d'une image requête avec les concepts de l'ontologie de domaine proposée. Ce processus est illustré comme suit :

- 1) Extraction des informations de l'image de télédétection :
 - a. Extraction des régions de l'image
 - b. Extraction des primitives visuelles de chaque région
 - c. Extraction des relations spatiales entre les régions
- 2) Modélisation floue des régions extraites
- 3) Matching floue : Détermination du concept (de l'ontologie) approprié pour chaque région de l'image.
- 4) Etablissement de l'annotation sémantique

L'annotation obtenue modélise les régions de l'image associées au concept correspondant de l'ontologie, les relations spatiales entre les objets classés, et les informations générales de l'image à savoir la date d'acquisition, le capteur, la

résolution, la dimension et le géo-référencement. Cette annotation constitue le « cas » sur lequel se base le processus de raisonnement à base de cas et d'ontologie.

3.2.2. Fusion floue d'annotations

Ce module est appliqué dans le cas d'une image multi-capteur représentant une seule scène. Dans ce cas, le processus d'annotation génère des annotations en fonction des capteurs de la scène. Afin de profiter de cette richesse d'information (vu la diversité des capteurs) et d'avoir une information fidèle à la réalité, nous proposons de fusionner ces annotations afin d'obtenir une plus précise, moins redondante et plus proche de la scène réelle. La fusion à ce stade se base sur les objets et les attributs de l'image. Le processus de fusion est basé sur l'opérateur d'union concernant les objets figurant dans l'une des annotations, l'opérateur d'intersection concernant les objets équivalents, et la fusion floue concernant les objets ayant des conflits (deux objets représentent la même région mais qui ont des interprétations différentes). Le résultat de ce module est une nouvelle annotation multi-capteurs.

3.2.3. Recherche d'annotations similaires

Après avoir modélisé l'image requête par des annotations représentant son contenu, nous proposons de procéder à une phase de recherche de cas similaires à partir d'une base d'images déjà annotées. L'objectif de cette phase est de retrouver les annotations similaires à celui de la requête annotée afin de profiter des situations passées. La recherche s'effectue en mesurant la similarité entre les annotations de la requête et du cas plutôt que les images en soi, en se basant sur (1) des méthodes d'appariement entre les objets et les relations spatiales, et (2) l'ontologie de domaine proposée. Le résultat de cette phase est un ensemble d'annotations similaires à la l'annotation de l'image requête ordonnées par le score de similarité.

3.2.4. Raisonnement et interprétation

Ce module exploite les résultats du module précédant et permet de déduire la susceptibilité d'une requête image à un risque naturel donné, en utilisant les atouts de l'ontologie en raisonnement inférentiel et le raisonnement à base de cas.

3.3. Application à l'étude de susceptibilité aux risques naturels

Il est bien connu que la dégradation et la disparition des couvertures végétales déclenchent et accélèrent les processus érosifs partout dans le monde, et plus particulièrement en Tunisie, sous l'action combinée de plusieurs facteurs à savoir le climat, la topographie, le couvert végétal et l'homme, entraînant de sérieux problèmes, tels que la réduction des terres arables, la diminution de la production agricole, la pollution des rivières, et la sécheresse. Nous avons appliqué notre méthodologie pour traiter ce problème afin de prévenir les zones à risque contre l'érosion en Tunisie et plus particulièrement à la région du Kef.

4. Organisation de la thèse

Dans une introduction générale, nous présentons le contexte général de la thèse, les problématiques ainsi que nos principales contributions.

Le manuscrit est ensuite structuré en trois parties : la première partie est un état de l'art sur les domaines traités dans notre thèse. Elle comporte les deux chapitres suivants :

Le chapitre 1 (*Annotation sémantique d'images de télédétection*) présente les images de télédétection comme la principale source de l'information géographique, leur mode d'acquisition et les prétraitements nécessaires pour leur exploitation. La deuxième section aborde l'annotation sémantique de l'image d'une façon générale, et présente quelques systèmes existants.

Le chapitre 2 (*Représentation des connaissances de l'image de télédétection par ontologie*) présente en détail la notion d'ontologie dans une première section. La deuxième section dresse un état de l'art sur les ontologies existantes modélisant l'image, suivie d'une synthèse et une discussion.

La deuxième partie décrit nos contributions à l'interprétation d'images de télédétection par ontologie. Elle comporte les deux chapitres suivants :

Le chapitre 3 (*Ontologie spectro-spatio-temporelle, floue, lourde et évolutive*) présente la première contribution concernant la proposition d'une ontologie répondant aux spécificités des images de télédétection. La première section présente le processus de

développement de l'ontologie proposée en se basant sur (1) l'enrichissement d'une ontologie de noyau, et (2) la réutilisation des ressources ontologiques existantes. La deuxième section, présentent l'ontologie résultante ainsi que sa définition formelle.

Le chapitre 4 (*Méthodologie d'annotation sémantique et l'interprétation de l'image de télédétection*) présente le cadre méthodologique du travail. La première section présente en détail l'architecture générale de la méthodologie qui se base essentiellement sur l'ontologie développée. La méthodologie proposée permet de générer une annotation sémantique floue de l'image de télédétection, et vise à déterminer la susceptibilité d'une scène à un risque naturel donné. Les sections suivantes présentent et illustrent chaque module de la méthodologie : (1) Annotation sémantique floue, (2) Fusion floue des annotations, (3) Recherche d'annotations, et (4) Raisonnement et interprétation.

La troisième partie présente le cadre applicatif de la méthodologie proposée. Nous avons choisi comme thème l'étude de la susceptibilité de scènes aux risques naturels, et plus précisément au risque d'érosion. Le chapitre 5 présente, dans la première section, les données et les outils utilisés pour valider l'expérimentation, la deuxième section illustre le processus d'annotation et d'interprétation sémantique d'une scène requête. Les résultats ainsi obtenus seront comparés aux données issues de l'expert de domaine, dans le but de démontrer l'efficacité de notre méthodologie en termes de précision.

Enfin, nous concluons en présentant un bilan général de l'ensemble de nos contributions et en évoquant de nouvelles perspectives de recherche futures.

Partie A

Etat de l'art

Annotation sémantique d'images de télédétection

L'objet de ce second chapitre est de présenter, dans une première partie, les images de télédétection, comme étant la principale source de l'information géographique, leur mode d'acquisition et les prétraitements nécessaires pour leur exploitation. La deuxième section aborde l'annotation sémantique de l'image d'une façon générale, et présente quelques systèmes existants.

1. L'image de télédétection

1.1. Définition d'une image de télédétection

L'image de télédétection constitue une source d'informations primordiales et précises sur l'occupation du sol et sur les phénomènes qui se déroulent à la surface de la planète. Elle est considérée comme la représentation graphique des mesures relevées par des capteurs placés en orbite autour de la terre (Bonn et al., 1992). Elle est constituée par un ensemble de pixels qui représentent la valeur de l'intensité du rayonnement réfléchi par l'objet observé dans la gamme de longueur d'ondes, auxquelles le capteur est sensible (Girard et Girard, 1999).

1.2. Principe d'acquisition : la télédétection

La télédétection est la principale méthode d'acquisition de l'image d'observation de la terre, elle est définie comme «*l'ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci.*»¹

Un système de télédétection fait intervenir quatre éléments de base (Goetz 1992) :

- Une source du rayonnement électromagnétique, qui peut être l'éclairement solaire dans le cas de la télédétection optique passive, ou le radar dans la télédétection passive.
- Une scène qui comprend la surface terrestre et l'atmosphère avec lesquelles interagit le rayonnement électromagnétique.
- Un capteur qui « recueille de l'énergie radiative provenant de la scène visée et délivre un signal électrique correspondant et mesurable » (Bonn et al., 1992).
- Un système d'interprétation et d'analyse de données permettant l'extraction d'informations utiles à partir des données issues du capteur et d'éventuelles données extérieures.

Pour plus d'informations sur les bases physiques de la télédétection, on se référera utilement aux manuels de Bonn et Rochon (1993) ou de Girard et Girard (1999).

1.3. Les capteurs d'acquisition

En télédétection, les capteurs permettent de mesurer des objets étudiés dans une bande donnée de longueur d'onde (Girard et Girard, 1999). Généralement, les capteurs sont de deux types : Les capteurs *actifs* qui envoient une énergie et recueillent la fraction que les objets en renvoient, et les capteurs *passifs* qui reçoivent l'énergie que les objets émettent ou réfléchissent.

Le capteur est caractérisé par sa résolution spatiale, spectrale, radiométrique et temporelle² :

- **La résolution spatiale** correspondant à la surface élémentaire d'échantillonnage observée instantanément par le capteur. Cette surface élémentaire correspond au pixel (*picture element*).

¹ Commission interministérielle de terminologie de la télédétection aérospatiale, 1988

² <http://www.rncan.gc.ca/>

- *La résolution spectrale* décrit la capacité d'un capteur à mesurer la réflectance dans des fenêtres de longueur d'onde comprises généralement entre 0.1 et 0.01 μm .
- *La résolution temporelle* ou répétitivité correspond à la période entre deux acquisitions de la même scène. Cette résolution ne dépend pas du capteur mais de l'orbite et du mode de manœuvre du satellite.
- *La résolution radiométrique* exprime l'aptitude du capteur, dans une bande spectrale considérée, à distinguer des signaux électromagnétiques d'énergies différentes. En effet, l'énergie mesurée par le capteur est codée en niveau de gris, et plus la gamme des niveaux de gris utilisée est étendue, plus la résolution radiométrique est grande. La plupart des images sont codées sur un octet (soit 256 niveaux de gris).

Le tableau 2.1 présente quelques capteurs et ses caractéristiques³.

	Landsat-5	SPOT-4	ERS-2
Date de lancement	1984 (USA)	1998 (France)	1995 (Europe)
Altitude	705 km	832 km	782-785
Capteurs et résolution spectrale	<i>Thematic Mapper</i> 0.45-0.52 μm TM1 0.52-0.60 μm TM2 0.63-0.69 μm TM3 0.76-0.90 μm TM4 1.55-1.75 μm TM5 10.4-12.5 μm TM6 2.10-2.35 μm TM7	<i>HRV (High resolution visible)</i> 0.50-0.59 μm XS1 0.61-0.69 μm XS2 0.79-0.90 μm XS3 0.51-0.73 μm P 1.5-1.75 μm MIR	<i>Altimètre radar</i> <i>Diffusomètre vent</i> <i>Sondeur hyperfréquences</i> <i>Capteur Infrarouge</i>
Résolution spatiale	30 m	10m (P) ; 20m (XS)	18m (SAR)
Répétitivité	16 jours	26 jours	35 jours

Tableau 1.1 : *Caractéristiques de quelques capteurs et satellites d'observation de la terre*

1.4. Types d'images de télédétection

1.4.1. L'image optique

Une image optique est acquise par un capteur passif. Selon le mode d'acquisition, il existe trois types d'images optiques : l'image panchromatique, l'image multi-spectrale et l'image hyper-spectrale (Bonn et al., 1992 ; Girard et al., 1999 :

³ <http://www.rncan.gc.ca/>

L'image panchromatique est obtenue à partir de l'enregistrement du rayonnement dans une seule bande électromagnétique. Elle permet d'obtenir des informations de type "géométrique" comme la forme, le périmètre et la surface.

L'image multi-spectrale est obtenue par des enregistrements simultanés d'un petit nombre de bandes spectrales. Ce type d'images permet d'obtenir des informations à caractère "qualitatif" (ex. classification des types de végétation sur des parcelles).

L'image hyper-spectrale est obtenue par des enregistrements simultanés d'un grand nombre de bandes spectrales étroites et contiguës. Les données hyper-spectrales fournissent donc une information plus détaillée, des propriétés spectrales d'une scène et permettent une identification et une discrimination plus précise des objets que les capteurs à larges bandes.

1.4.2. *L'image Radar*

L'image Radar est acquise par un capteur Radar (Radio Detection And Ranging). Elle complète de manière efficace les images optiques, et permet, en effet, le suivi de vastes régions indépendamment des conditions climatiques et météorologiques, et est particulièrement sensible à la topographie, à la rugosité et à l'humidité de la surface.

1.5. *Prétraitement des images de télédétection*

Les images satellitales acquises par les satellites ne sont pas directement exploitables, elles nécessitent un ensemble de traitements préalables pour qu'elles soient interprétables. Ces opérations de prétraitement se divisent en corrections géométriques, en corrections radiométriques et en corrections atmosphériques (Bonn et al., 1992 ; Girard et al., 1999).

1.6. *Spécificités des images de télédétection*

1.6.1. *L'image de télédétection, information liée au temps et au capteur*

L'interaction rayonnement-matière est une des bases de l'interprétation des images de télédétection. Un objet, à un moment donné, dans une position géographique donnée, vu sous un angle donné et recevant un rayonnement donné, présente un comportement spectral qui lui est propre (Girard et al., 99). Le comportement spectral d'un objet varie suivant le temps, le lieu, le mode de saisie de l'information et du rayonnement incident (Figure 1.1).

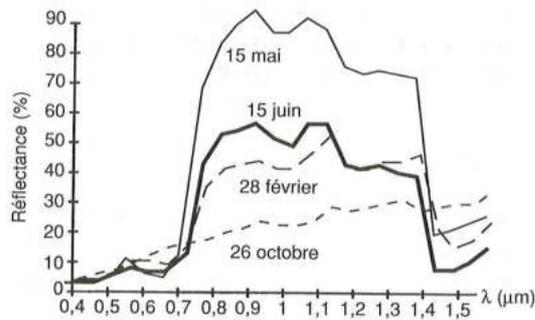


Figure 1.1 : Mesures de réflectance à différentes dates pour une parcelle de blé d'hiver (Source : Girard et al., 1999)

Les objets étudiés sont multiples, mais peuvent se ramener très souvent à quelques cas généraux : la végétation (et ce que ressort de l'organique), les sols (et ce qui ressort de minéral), l'eau, la neige, et la glace. La figure 1.2 illustre le comportement spectral de quelques matériaux suivant les capteurs SPOT et LandSat.

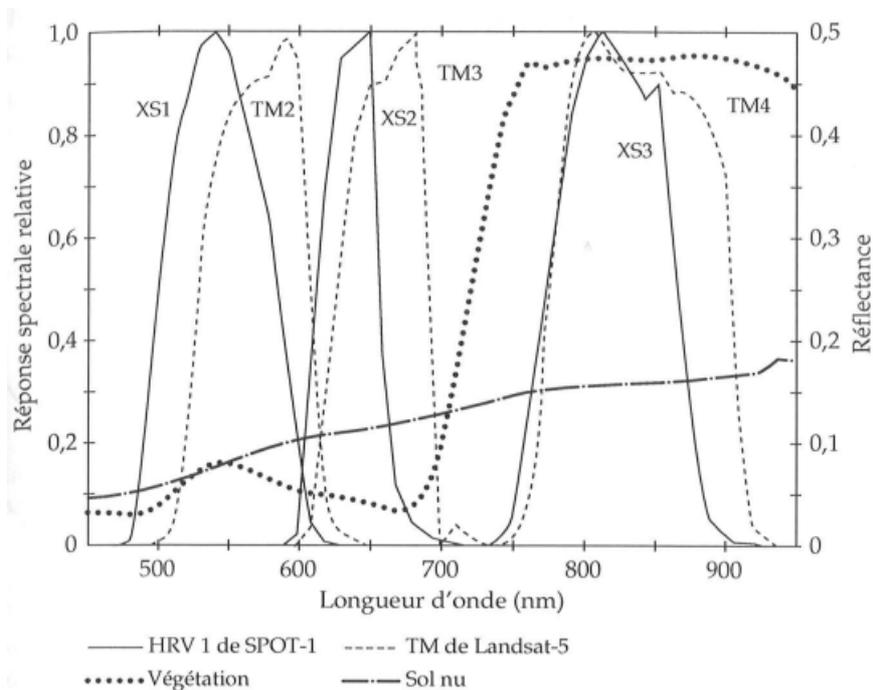


Figure 1.2 : Comportements spectraux de quelques matériaux suivant les capteurs SPOT et LandSat (Source : Girard et al., 1999)

1.6.2. *Imperfections entachant l'image de télédétection*

Les images de télédétection présente plusieurs types d'imperfections provenant de sources diverses à savoir :

- Les erreurs de mesures des capteurs qui constituent la source de base des images satellites.
- Les limites des méthodes de compression et de réduction de données.
- La perte d'information due aux limites des méthodes de traitement et d'extraction d'informations.

Une bonne interprétation d'images satellite doit tenir compte de l'imperfection entachant l'image de télédétection.

1.7. *Interprétations des images satellitales*

L'interprétation des images satellitales consiste à donner du sens aux données provenant de la scène analysée, c'est-à-dire à reconnaître les différentes entités qui composent l'image, à comprendre leur organisation spatiale et à construire une description de la scène. L'interprétation d'image fait donc appel à un certain nombre de connaissances *a priori* sur :

- ✓ la scène et sur les objets qu'elle contient (modélisation des objets et de leurs relations)
- ✓ le domaine d'application (expertise).
- ✓ les conditions d'acquisition des images.

L'interprétation d'images se base alors sur des :

- ✓ Modèles et formalismes pour représenter les connaissances.
- ✓ Modules d'apprentissages pour acquérir et structurer les connaissances.
- ✓ Opérateurs d'analyse d'image, paramétrables et adaptatifs.
- ✓ Opérateurs de reconnaissances ou de mise en correspondance entre les informations *a priori* et les résultats de l'analyse.

2. Annotation sémantique d'images guidée par ontologie

2.1. Annotation et problème du fossé sémantique

L'annotation d'images se trouve au point de convergence de deux domaines de recherche voisins : la reconnaissance d'objets et la recherche d'images par le contenu. Elle est définie comme le processus qui consiste à décrire et caractériser le contenu de l'image par des mots clés ou par des descripteurs visuels.

Dans le cas d'annotation par mots clés, l'image est décrite manuellement par des mots du langage naturel par des humains. Cependant, cette technique présente quelques limites liées aux choix de termes d'annotations. En effet, un mot annotant une image peut comporter plusieurs significations. Par exemple, "avocat" peut désigner à la fois un arbre ou un métier. De plus, Le contenu d'une image peut être décrit par un ensemble de mots plus ou moins précis : "rivière", "étendue d'eau", ou "parcours d'eau" sont différentes manières de décrire un même objet dans une image. En outre, les personnes qui annotent les images vont considérer un niveau de spécification qui correspond à leur objectif ou à leur degré de connaissance.

Dans le cas de l'annotation de l'image par des descripteurs visuels (ou annotation automatique), l'image est décrite à travers des caractéristiques de bas niveau extraites automatiquement à savoir la couleur, la texture et la forme. Ces caractéristiques représentent soit la totalité de l'image ou des régions de celle-ci. Il existe de très nombreux descripteurs visuels dans la littérature. On pourra en trouver un bon état de l'art dans (Smeulders et al., 2000).

Ce type d'annotation présente des ambiguïtés au niveau de la correspondance (ou ancrage) des données numériques extraites de l'image avec leurs interprétations sémantiques. Ce problème s'appelle *fossé sémantique* entre l'information de bas niveau et celle de haut niveau. Smeulders et al. (2000) le définit comme «*le manque de concordance entre l'information que l'on peut extraire des données visuelles et l'interprétation que les mêmes données ont pour un utilisateur dans une situation donnée*».

Récemment, plusieurs travaux se sont intéressés à l'utilisation des avancées scientifiques dans le domaine de l'intelligence artificielle pour franchir le problème de fossé sémantique. Ces systèmes se basent sur les connaissances de domaine représentées par des ontologies. Ces dernières offrent une compréhension commune, structurée et partagée d'un domaine ou d'une tâche, qui peut être utilisée pour la communication entre les personnes et les machines.

2.2. Systèmes d'annotation existants

Nous présentons dans la section suivante un survol sur quelques systèmes d'annotations d'images basées sur les ontologies.

Le travail décrit dans Mezaris et al. (2003) repose sur la définition d'ontologies spécifiques aux objets représentés dans les images (Object Ontology). L'ontologie développée représente les caractéristiques visuelles des objets de l'image par une description symbolique permettant une interprétation de haut niveau (figure 1.3). Les auteurs ont appliqué des algorithmes de segmentation non supervisée pour extraire les régions de l'image, et d'autres algorithmes permettant le passage entre une description symbolique et une analyse de bas niveau de l'image.

Les caractéristiques modélisées sont :

- La luminosité (exprimée par cinq paramètres) et la distribution des couleurs dans la gamme vert-rouge et bleu-jaune (sept paramètres).
- La position de la région dans l'image : horizontale et verticale (trois paramètres par orientation).
- La taille de l'objet (trois paramètres)
- La forme de l'objet (trois paramètres)

Les régions de l'image supposées contenir l'objet sont ainsi définies par une série de caractéristiques de bas niveau dans une forme compacte.

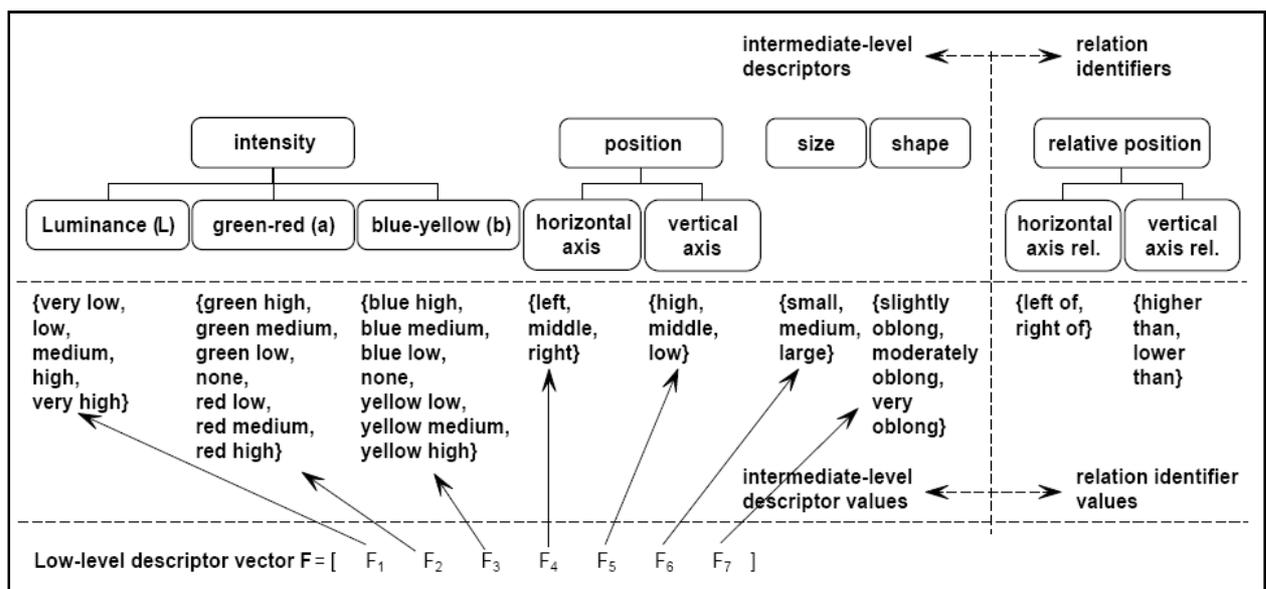


Figure 1.3 : Ontologie d'objets

Le travail décrit dans (Breen et al., 2002) consiste à reconnaître et annoter des images de ballons en utilisant une ontologie de domaine de sport. La reconnaissance s'effectue en se basant sur un réseau de neurones artificiel permettant d'apprendre les concepts de l'ontologie, et ensuite de classifier le type de ballon entré en input (ballon de foot, ballon de tennis, ballon de basket...).

Ruan et al. (2006) présentent une plate-forme de recherche d'images satellitaires en utilisant des ontologies pour annoter les images à partir de ses caractéristiques de bas niveau. Le système utilise un raisonneur pour rechercher les concepts de l'ontologie correspondants aux requêtes qui peuvent être textuelles, images ou région d'une image. Les auteurs ont utilisé seulement trois relations spatiales : *est-un*, *instance-de* et *partie-de*.

Hudelot et al., (2005) proposent un système générique d'interprétation sémantique des images de pathologie végétale en se basant sur une ontologie de concepts visuels et une ontologie de traitement d'images (figure 1.4). L'ontologie de concepts visuels est une hiérarchie de concepts qui s'organise en :

- concepts spatio-temporels
- concepts de couleur (teinte, bleu, clair,...)
- concepts de texture (texture régulière, texture maillée,...)

L'ontologie de traitement d'image regroupe les concepts suivants :

- concepts d'entités image : ils représentent les différentes structures de données qui peuvent être extraites des images (région, contour,...)
- concepts de descripteurs d'images: ils représentent les différentes caractéristiques pouvant être mesurées dans l'image
- concepts de fonctionnalités de traitement d'images : ils représentent des buts génériques de traitement d'images (segmentation, extraction de caractéristiques, ...)

La correspondance entre les données images et les concepts visuels se fait à travers des algorithmes d'apprentissage.

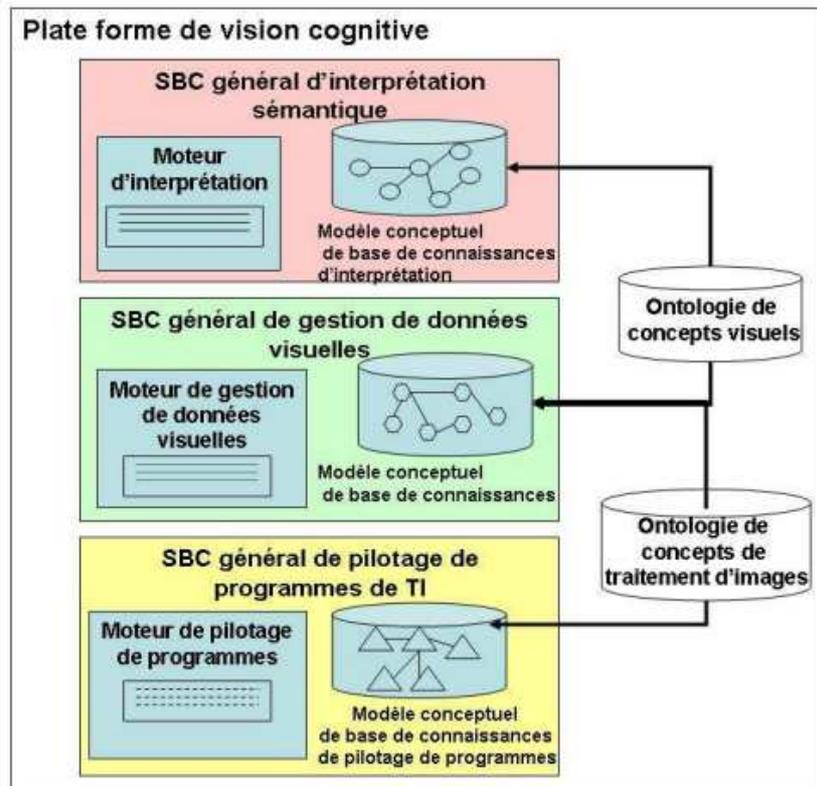


Figure 1.4 : Architecture de la plate-forme d'interprétation d'image
(Source : Hudelot et al., 2005)

Les auteurs dans (Wang et al., 06) ont proposé une ontologie intégrant des informations de bas niveau et de haut niveau de l'image (figure 1.5). L'ontologie développée part d'une ontologie de description basée sur la ressource sémantique « BBC Science and Nature Animal Category » contenant 620 concepts du domaine animal. Cette ontologie est enrichie par la suite par deux autres ontologies, la première décrit les caractéristiques visuelles extraites à partir des images (bas niveau), et la deuxième provenant des ressources textuelles (haut niveau) extraites à partir des textes décrivant les images. Elle est appliquée dans la recherche des documents multimédia contenant à la fois du texte et des images.

Dans le domaine de télédétection, Durand et al. (2007), proposent un système de reconnaissance d'objets d'une image satellitale en se basant sur une ontologie développée par des experts du domaine urbain (figure 1.6). La reconnaissance d'objets se fait en effectuant la correspondance entre les objets de l'image avec les concepts de l'ontologie. L'algorithme proposé consiste à rechercher l'objet extrait de l'image dans l'ontologie en comparant ses descripteurs bas niveau avec ceux des concepts de l'ontologie, et d'établir un score de similarité avec chaque concept de l'ontologie.

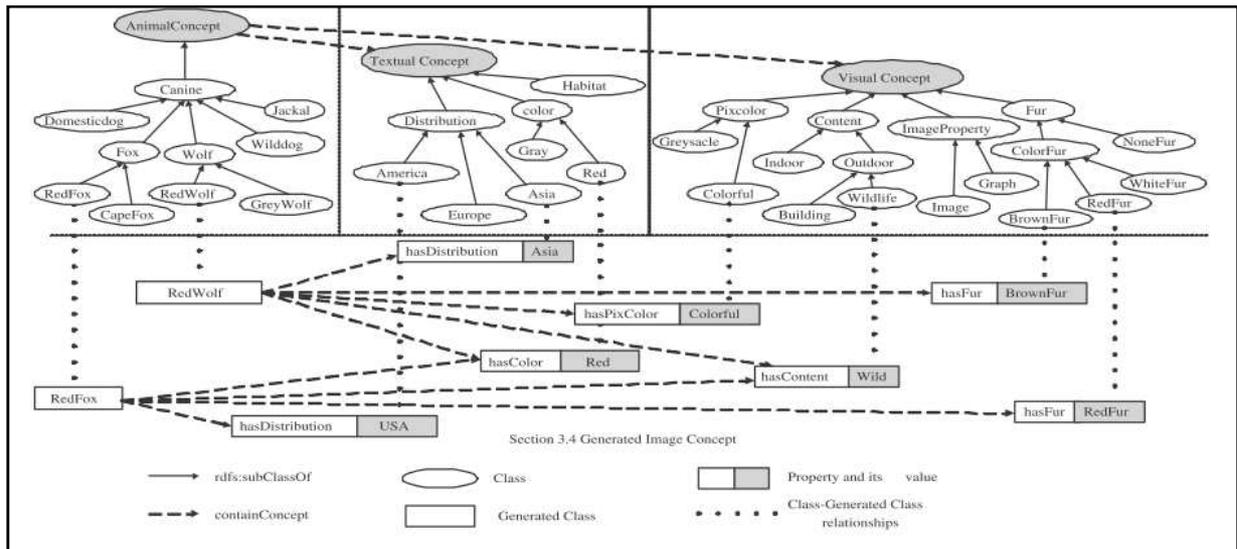


Figure 1.5 : *Ontologie regroupant les concepts visuels et textuels*

Mechouche et al. (2009), proposent un système hybride combinant des connaissances numériques et symboliques pour l'annotation sémantique des structures anatomiques tridimensionnelles, extraites automatiquement des images de résonance magnétique du cerveau. Dans ce cas, les connaissances numériques consistent en un atlas des gyri du cerveau, et les connaissances symboliques consistent en une ontologie formelle de l'anatomie du cortex cérébral. Ceci est fait en réutilisant une ontologie de haut niveau existante concernant l'annotation des régions d'intérêt en neuro-imagerie. Le processus d'annotation consiste à générer des labels modélisant les structures anatomiques du cerveau.

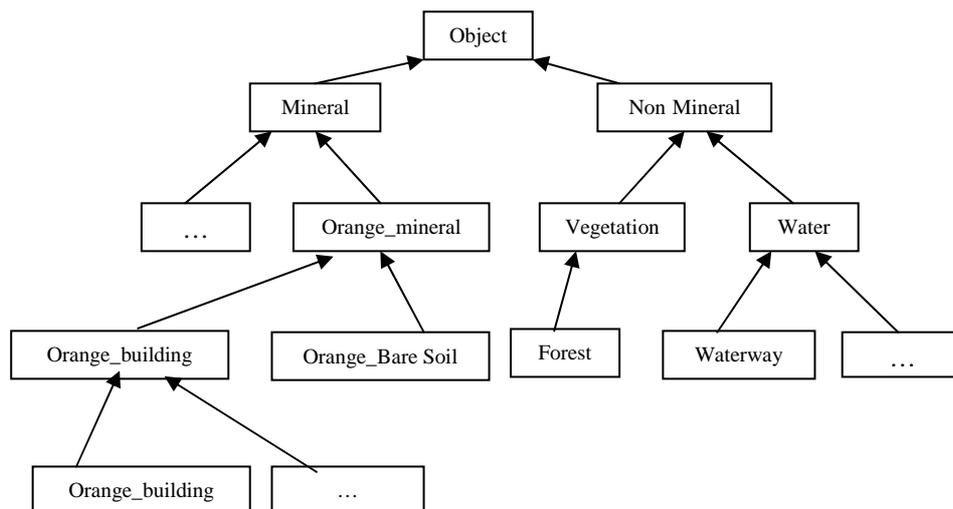


Figure 1.6 : *Extrait de l'ontologie de Durant et al., (2007)*

2.3. Synthèse et discussion

2.3.1. Rôle de l'ontologie dans l'annotation sémantique d'images

Les travaux cités ci-dessus ont montré que l'usage de l'ontologie dans un système d'annotation sémantique permet de :

- Représenter l'information visuelle ;
- Réduire le fossé sémantique entre les primitives de l'image et leur description sémantique ;
- Reasonner et inférer des nouvelles connaissances ;

2.3.2. Composants d'un système sémantique d'annotation d'images par ontologie

L'annotation basée sur ontologie dépend essentiellement de deux volets :

- 1) L'ontologie à utiliser pour guider l'annotation :
 - Domaine d'application ;
 - Type de l'ontologie ;
 - Méthode de construction ;
 - Niveau de complexité (légère/lourde) ;
- 2) Le processus d'annotation : La mise en correspondance entre les données extraites de l'image et les concepts de l'ontologie :
 - Type d'annotation : manuelle/automatique/semi-automatique ;
 - Apprentissage adopté ;

Conclusion

Dans ce chapitre, nous avons présenté les spécificités des images de télédétection. Nous avons mis l'accent sur l'annotation sémantique de l'image de télédétection. Nous avons fait ressortir ainsi que l'utilisation des ontologies s'avère très intéressante dans la représentation de connaissances au niveau de l'image.

Représentation des connaissances de l'image de télédétection par ontologie

Dans ce chapitre, nous abordons, dans une première section, la notion d'ontologie. Nous présentons, dans une deuxième section, un état de l'art sur les ontologies liées au domaine de la géographie et de la télédétection suivi d'une synthèse et une discussion.

1. La notion d'ontologie

Le terme "Ontologie" est un terme philosophique qui signifie « *Partie de la métaphysique qui s'applique à l'être en tant qu'être, indépendamment de ses déterminations particulières* » (Le Petit Robert). Avec l'émergence de l'ingénierie des connaissances (IC), l'ontologie est introduite en Intelligence Artificielle (IA) comme réponse aux problématiques de représentation et de manipulation des connaissances au sein des systèmes informatiques.

1.1. Définitions

Il est difficile d'attribuer à la notion d'ontologie une unique définition définitive puisque celle-ci a été employée dans des contextes différents. Neeches et al., (1991) furent les premiers à proposer une définition de l'ontologie à savoir : « *une ontologie définit les termes et les relations de base du vocabulaire d'un domaine ainsi que les règles qui indiquent comment combiner les termes et les relations de façon à pouvoir étendre le vocabulaire* »

Gruber (1993) a défini l'ontologie comme « *une spécification explicite d'une conceptualisation* ». Cette définition a été modifiée légèrement par Borst (1997) comme « *spécification formelle d'une conceptualisation partagée* ».

Ces deux définitions ont été combinées par Studer (1998) comme « *spécification formelle et explicite d'une conceptualisation partagée* » :

- « **Conceptualisation** » fait référence à un modèle abstrait de certains phénomènes du monde, modèle qui identifie les concepts pertinents de ce phénomène.
- « **Explicite** » signifie que le type des concepts utilisés et les contraintes sur leur utilisation sont explicitement définis.
- « **Formelle** » se réfère au fait que l'ontologie doit être compréhensible par les machines.
- « **Partagée** » reflète la notion de connaissance consensuelle décrite par l'ontologie, c'est-à-dire qu'elle n'est pas restreinte au point de vue de certains individus seulement, mais reflète un point de vue plus général, partagé et accepté par un groupe.

Selon Swartout et al., (1997), une ontologie est « *un ensemble de termes structurés de façon hiérarchique, conçue afin de décrire un domaine et qui peut servir de charpente à une base de connaissances* ».

Sowa (1995) a mentionné que le but d'une ontologie est l'étude des catégories de concepts qui existent ou peuvent exister dans certains domaines. Le résultat de cette étude, appelé une ontologie, est un catalogue de types de choses qui existent dans un domaine d'intérêt dans la perspective d'utiliser un langage formel pour parler de domaine.

On remarque que les définitions, malgré leur diversité, offrent des points de vue complémentaires. Ainsi, une ontologie offre les moyens pour présenter les concepts d'un domaine en les organisant hiérarchiquement et en définissant leurs propriétés sémantiques dans un langage formel de représentation des connaissances.

1.2. Constituants de l'ontologie

Une ontologie est constituée de plusieurs composants à savoir : les concepts souvent représentés par des termes, les relations entre ces concepts, les axiomes et les instances.

1.2.1. Les concepts

Selon Uschold et al., (1996), un concept représente un objet matériel, une notion ou une idée. Il est composé de trois parties : un ou plusieurs termes, une notion et un ensemble d'objets :

- **Le terme**, appelé aussi *label*, permet de désigner le concept.
- **La notion**, également appelée intension du concept, correspond à la sémantique du concept, exprimée en termes de propriétés et d'attributs.
- **L'ensemble d'objets**, également appelé extension du concept, regroupe les objets manipulés à travers le concept ; Ces objets sont appelés instances du concept.

Guarino (1997) a identifié les propriétés que peut avoir un concept de l'ontologie :

- **La généricité** : un concept est générique s'il ne possède pas d'extension.
- **L'identité** : un concept porte une propriété d'identité si cette propriété permet de conclure quant à l'identité de deux instances de ce concept. Cette propriété peut porter sur des attributs du concept ou sur d'autres concepts.
- **La rigidité** : un concept est dit rigide si toute instance de ce concept en reste instance dans tous les mondes possibles.
- **L'anti-rigidité** : un concept est anti-rigide si toute instance de ce concept est essentiellement définie par son appartenance à l'extension d'un autre concept.
- **L'unité** : un concept est un concept unité, si pour chacune de ses instances, les différentes parties de l'instance sont liées par une relation qui ne lie pas d'autres instances du concept.

De même, deux concepts peuvent avoir les propriétés suivantes :

- **L'équivalence** : deux concepts sont équivalents s'ils ont la même extension.

- **La disjonction** : deux concepts sont disjoints si leurs extensions sont disjointes.
- **La dépendance** : Un concept C_1 est dépendant d'un concept C_2 si pour toute instance de C_1 il existe une instance de C_2 qui ne soit ni partie ni constituant de l'instance de C_1 . Exemple : parent est un concept dépendant de enfant (et vice-versa).

1.2.2. *Les relations*

Les relations représentent un type d'interaction entre les concepts d'un domaine. Elles se décomposent généralement en deux catégories : les relations *taxonomiques* et les relations *associatives ou non taxonomiques*.

Les relations taxonomiques organisent l'ensemble des concepts de l'ontologie dans une structure arborescente. Selon Guarino (2000), un concept C_1 subsume un concept C_2 si toute propriété sémantique de C_1 est aussi une propriété sémantique de C_2 , autrement dit, C_1 est plus spécifique que C_2 .

Les relations associatives sont des relations d'interaction entre deux concepts qui sont non taxonomiques. « Associé-a », « est-situé-dans », « adjacent-à », « déconnecté-à » sont des exemples de relations non taxonomiques.

1.2.3. *Les axiomes*

Les axiomes ont pour but de définir dans un langage logique la description des concepts et des relations permettant de représenter leur sémantique. Ils représentent les intentions des concepts et des relations du domaine et, de manière générale, les connaissances n'ayant pas un caractère strictement terminologique (Staab, 2000).

L'intégration des axiomes dans une ontologie peut avoir plusieurs objectifs :

- Définir la signification des composants,
- Définir des restrictions sur la valeur des attributs,
- Définir les arguments d'une relation,

- Vérifier la validité des informations spécifiées ou en déduire de nouvelles.

1.2.4. Les instances

Les instances constituent la définition extensionnelle de l'ontologie, et véhiculent les connaissances (statiques, factuelles) à propos du domaine du problème.

1.3. Classification des ontologies

Les ontologies peuvent être classifiées selon plusieurs critères : l'objet conceptualisation, l'engagement sémantique et le niveau de détail.

1.3.1. Classification selon l'objet de conceptualisation

Cette classification est basée sur le contenu de connaissance que l'ontologie représente (Guarino, 1997 ; Grüninger et al., 1995 ; Uschold et al., 1996 ; Sowa, 1995) :

Ontologie de haut niveau : appelée aussi ontologie abstraite, supérieure ou fondationnelle, cette ontologie porte les concepts les plus généraux tels que l'espace, le temps, la matière, les objets, les événements, et les actions qui sont indépendants d'un domaine particulier. Elle spécifie un formalisme de description qui fournit une structure de représentation et des primitives pour décrire les concepts des ontologies de domaine et des ontologies génériques, et permet de définir des concepts abstraits et peuvent être réutilisées pour définir des concepts spécifiques.

Ontologie générique : appelée aussi méta-ontologie ou noyau d'ontologie, elle véhicule des connaissances génériques qui doivent être assez générales pour être réutilisées dans différents domaines. Elle décrit des concepts généraux, indépendants d'un domaine ou d'un problème particulier.

Ontologie de domaine : Il s'agit d'une conceptualisation spécifique à un domaine particulier. Elle rend compte du vocabulaire d'un domaine spécifique au travers de concepts et de relations qui modélisent les principales activités, les théories et les principes de base du domaine en question. Les concepts d'une

ontologie de domaine sont souvent définis comme une spécialisation des concepts des ontologies génériques.

Ontologie de tâche : Elle utilisée pour conceptualiser des tâches spécifiques dans les systèmes, telles que les tâches de diagnostic, de planification, de conception, de configuration, de tutorat, soit tout ce qui concerne la résolution de problèmes. Elle régit un ensemble de vocabulaires et de concepts qui décrit une structure de résolution des problèmes inhérente aux tâches et indépendante du domaine. Selon (Grüninger et al., 95), l'ontologie de tâche caractérise l'architecture computationnelle d'un système à base de connaissances qui réalise une tâche.

Ontologie d'application : Elle décrit des concepts qui dépendent à la fois d'un domaine particulier et d'une tâche particulière. Elle est souvent une spécialisation à la fois des ontologies de domaine et des ontologies de tâches et correspond aux rôles joués par les entités de domaine lorsqu'elles effectuent certaines activités. Ce sont les ontologies les plus spécifiques.

1.3.2. Classification selon l'engagement sémantique

Cette classification se base sur le niveau sémantique des connaissances que l'ontologie représente (ressource terminologique versus ressource conceptuelle).

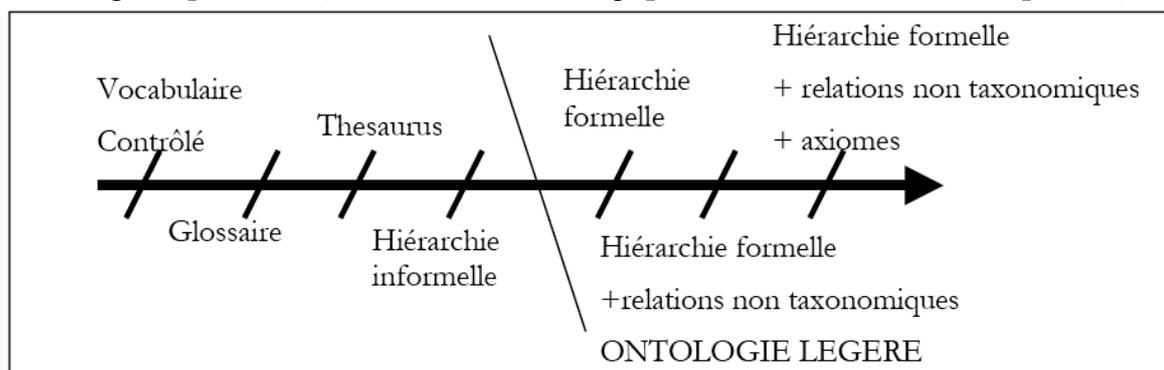


Figure 2.1 : Les différentes représentations de connaissance à partir de leur degré d'engagement sémantique (inspiré de (Lassila et al., 2001))

La figure 2.1 montre une échelle des degrés de formalisation et du niveau d'engagement sémantique en représentation de connaissances. Le degré d'engagement sémantique correspond au niveau de spécification formelle permettant de restreindre l'interprétation de chaque concept et ainsi d'en donner la sémantique (Bachimont et al., 2000).

Cette échelle va du vocabulaire contrôlé qui est un simple ensemble de mots appartenant à un domaine, aux ontologies lourdes (heavy-weight ontologies) comprenant les termes désignant les concepts et relations du domaine, leurs propriétés conceptuelles, et toutes les connaissances nécessaires à la description de la sémantique du domaine. Entre les deux, des glossaires aux ontologies légères, les représentations de connaissances offrent des possibilités croissantes d'intégrer la sémantique du domaine, en spécifiant toutefois toujours la terminologie.

1.3.3. Classification selon le niveau de détail

Par rapport au niveau détail, deux catégories d'ontologies peuvent être identifiées (Guarino, 1997):

Granularité fine : correspondant à des ontologies très détaillées, possédant ainsi un vocabulaire plus riche capable d'assurer une description détaillée des concepts pertinents d'un domaine ou d'une tâche.

Granularité large : correspondant à un vocabulaire moins détaillé c'est le cas des ontologies de haut niveau possèdent une granularité large, compte tenu que les concepts qu'elles traduisent sont normalement raffinés dans d'autres ontologies de domaine ou d'application.

1.4. Méthodologies de construction des ontologies

Plusieurs méthodologies ont été proposées pour la construction d'une ontologie. Dans ce qui suit, nous passons en revue les méthodes les plus citées dans la littérature.

1.4.1. La méthode "Enterprise Ontology"

Uschold et King (1995) ont proposé la première méthode de construction d'ontologies qui est fondée sur l'expérience réussie lors du développement de l'ontologie d'Entreprise.

Cette méthode repose sur les étapes suivantes :

- Identification du pourquoi de l'ontologie ;
- Construction de l'ontologie (identification des concepts et relations clés) et intégration d'ontologies existantes (comment les utiliser) ;
- Evaluation et documentation de l'ontologie.

1.4.2. La méthodologie On-To-Knowledge (OTK)

La fin du projet On-To-Knowledge (Staab, 2001) est d'appliquer les ontologies aux informations disponibles électroniquement pour progresser la qualité de la gestion de la connaissance au sein des organisations. Dans ce cadre, l'idée était le développement d'une méthodologie et des outils pour l'accès intelligent aux sources de large volume.

Cette méthode repose sur les étapes suivantes :

- Etude de faisabilité par l'identification des problèmes ainsi que les solutions potentielles.
- Démarrage (Kickoff) : le résultat de cette étape est un document de spécifications des besoins d'ontologie, qui décrit le contexte et l'objectif de l'ontologie, ainsi les directives de conception (par exemple, les conventions de nommage), les sources de connaissance et les utilisateurs potentiels.
- Raffinement : l'objectif de cette étape est de produire une application orientée application conformément aux spécifications données à l'étape de Kickoff.
- Evaluation : qui consiste à contrôler si l'ontologie satisfait les besoins.
- La maintenance : qui consiste à préciser qui va s'occuper de la maintenance et comment.

1.4.3. La méthodologie METHONTOLOGY

Cette méthodologie a été développée à l'Université Polytechnique de Madrid. En utilisant cette méthode le développement d'ontologies est divisé en six tâches

- Spécification : déterminer l'utilisation future de l'ontologie ;
- Conceptualisation : obtenir un modèle du domaine au niveau des connaissances ;
- Formalisation : transformation du modèle conceptuel en modèle formel ;
- Intégration : réutilisation d'autres ontologies ;
- Implémentation : construction d'un modèle opératoire utilisable par un ordinateur.
- Maintenance : mise-à-jour de l'ontologie en cas de besoin.

1.4.4. La méthodologie STANFORD

Cette méthode a été développée à l'université de STANFORD (Natalya, 2002), qui comporte sept étapes qui sont les suivantes :

- Déterminer le domaine et la portée de l'ontologie ;
- Réutiliser des ontologies existantes ;
- Enumérer les termes importants de l'ontologie ;
- Définir les classes et la hiérarchie de classes ;

- Définir les propriétés des classes (les attributs) ;
- Définir les attributs ;
- Créer des instances des classes dans la hiérarchie ou exemples ;

Plusieurs outils peuvent être utilisés pour construire des ontologies en suivant cette méthodologie, par exemple, Protégé- 2000, OntoEdit, etc.

1.4.5. La méthodologie Kactus

Cette méthodologie vise à réutiliser des ontologies existantes et présentes des mécanismes permettant cette réutilisation. Ce principe est intéressant dans la mesure où il évite de construire une ontologie à partir de rien. Cette méthodologie repose sur trois étapes :

- Spécification de l'application : Cette phase consiste à limiter le contexte d'application.
- Conception préliminaire : basée sur des ontologies de haut niveau où la liste des termes et des tâches développées pendant la phase précédente est utilisée pour l'obtention de plusieurs visions du modèle globale conformément aux catégories du plus haut niveau ontologique.
- Raffinement et structuration de l'ontologie : Cette étape consiste à raffiner le travail fait dans les deux étapes précédentes afin d'assurer une bonne conception.

1.4.6. La méthodologie TOVE¹

La méthode de Grüninger et Fox (2005) est fondée sur l'expérience du développement de l'ontologie du projet TOVE (TOrento Virtual Enterprise). Elle aboutit à la construction d'un modèle logique de connaissances. L'ontologie est développée selon les étapes suivantes :

- Identification des problèmes dépendants d'une application ;
- Formulation de questions informelles auxquelles l'ontologie doit répondre.
- Spécification d'une terminologie à partir des termes apparaissant dans les questions.
- Spécification formelle des axiomes et des définitions pour les termes de la terminologie.
- Evaluation de la complétude de l'ontologie.

¹ <http://www.eil.utoronto.ca/tove/ontoTOC.html>

1.4.7. La méthodologie SENSUS

C'est une méthode qui consiste à construire une ontologie de domaine à partir d'une ontologie plus grande, l'ontologie SENSUS (Swartout, 1997). La méthode permet d'unir les termes caractéristiques du domaine à cette ontologie et d'éliminer les termes qui ne sont pas pertinent dans la nouvelle ontologie qu'on souhaite créer.

Cette méthodologie suit les étapes suivantes :

- Identifier les termes clés du domaine avec l'aide des experts ;
- Relier manuellement les termes clés à SENSUS ;
- Inclure tous les concepts qui se trouvent sur le chemin depuis le terme clé jusqu'à la racine de SENSUS ;
- Ajouter les nouveaux termes de domaine ;

1.5. Intégration et alignement d'ontologies

L'intégration d'ontologie est définie comme un processus qui permet de combiner des ontologies pour construire une nouvelle ontologie (Pinto et al., 1999). Les ontologies sources n'appartiennent pas au même domaine de discours. À cette fin, il est nécessaire d'établir des liens sémantiques entre les entités appartenant à deux ontologies différentes, par le biais d'un processus appelé alignement d'ontologies. Étant données deux ontologies, l'alignement produit un ensemble de correspondances, chacune liant deux entités par une relation (équivalence, subsomption, incompatibilité, etc.), éventuellement munie d'un degré de confiance (Euzenat et al., 2007). L'établissement de ces correspondances est fondé sur l'utilisation des mesures de similarités, qui peuvent être classées comme suit (Rahm et al., 2001, bach et al., 2006):

Les méthodes terminologiques : comparent les labels des entités. Elles sont décomposées en approches purement syntaxiques et lexicales.

Les méthodes de comparaison des structures internes : comparent les structures internes des entités (intervalle de valeur, cardinalité d'attributs, etc.).

Les méthodes de comparaison des structures externes : comparent les relations d'entités avec d'autres.

Les méthodes de comparaison des instances : comparent les extensions des entités, c'est à dire l'ensemble des autres entités qui lui sont attachées (instances et classes);

La méthode sémantique : compare les interprétations (ou plus exactement les modèles) des entités.

1.6. Langages de représentation des ontologies

Plusieurs langages ont été proposés pour la formalisation des ontologies. Les premiers langages sont issus de l'intelligence artificielle à savoir les réseaux sémantiques, les graphes conceptuels, les frames et les Logiques de Description. D'autres langages sont issus du Web sémantique tels que le langage XML, OIL, OWL, etc.

1.6.1. *Les réseaux sémantiques*

Un réseau sémantique est une représentation graphique d'une conceptualisation d'une (ou plusieurs) connaissance humaine (Quillian et al., 1968). Il est représenté sous la forme d'un graphe étiqueté et orienté. Les inférences possibles dépendent de la nature des liens. Cependant, ce type de définition ne concerne que la structure du graphe et ne permet pas d'ajouter de l'information sémantique. De nombreuses études ont montré que ce type de graphe manque de précision sémantique et mène à des confusions entre les relations et aussi entre les classes et individus. Elles ont mené à la définition de nouveaux formalismes tels que les frames, les logiques de description et les graphes conceptuels.

1.6.2. *Les frames*

Les frames (Minsky et al., 1975) sont présentés comme étant une structure de données capable de représenter des objets structurés. Un frame représente donc une classe ou un objet. Les frames sont organisés dans une hiérarchie suivant un lien de spécification. Les composants du frame sont appelés « slots ». Ils sont considérés comme des attributs de la structure. Ils peuvent être de plusieurs natures : valeur de l'attribut (qui peut être vide), ensemble de valeurs, restriction de valeurs, valeur par défaut, une propriété avec un autre frame, une combinaison des différents cas. L'intérêt des frames est qu'ils permettent de

représenter la façon de penser d'experts en fournissant une représentation structurée et concise des relations utiles. L'information peut être partagée entre plusieurs frames grâce à l'héritage.

1.6.3. Les logiques de description

Les logiques de description issues des frames reposent sur trois notions de base : les concepts représentant des classes (ensemble d'objets), les rôles (relations liant deux objets) et les individus (objets représentant les classes qu'ils instancient). Pour décrire ces éléments, deux structures sont utilisées : la T-BOX et la A-BOX. La T-BOX (boîte terminologique) comprend la description des concepts et des rôles. Cette description est structurée à l'aide du lien hiérarchique sorteDe. Deux concepts particuliers figurent au minimum dans la T-BOX : le concept le plus générique (anything) et le concept le plus spécifique (nothing). La A-BOX (boîte assertionnelle) est constituée des individus, de leur description et des règles qui leur sont attachés. Les inférences reposent sur la reconnaissance d'instances de concepts à partir de leur définition, la détection des concepts plus généraux ou plus spécifiques, et la classification ordonnant les concepts dans la hiérarchie. Les logiques de description sont plus flexibles que les frames et reposent sur une sémantique et une syntaxe rigoureuses (Baader et al., 1991).

1.6.4. Les graphes conceptuels

Les graphes conceptuels ont été présentés par Sowa (1984). Ils utilisent une notation à base de graphes. Ils ont été définis comme un langage pivot entre le langage naturel et la logique du premier ordre. Ils visent à formaliser les relations entre prédicats et arguments dans une phrase. Ils sont composés de deux types de nœuds étiquetés : les nœuds concepts et les nœuds relations. Les nœuds concepts et les nœuds relations sont respectivement typés par des types de nœuds et des types de relations, organisés suivant un ordre partiel. Les graphes conceptuels peuvent être vus comme des schémas permettant de représenter graphiquement des formules logiques, ou bien des schémas sans contraintes, servant juste d'interface « graphique » à la représentation de formules ou bien comme des graphes munis d'opérations de graphes permettant le raisonnement et leur manipulation en s'appuyant sur la théorie des graphes.

1.6.5. *Les langages du Web sémantique*

Plusieurs langages ont été proposés pour le Web sémantique dérivés du XML (eXtensible Markup language) qui permet de générer des balises pour la structuration de données et de documents (Bradley, 2001). Le langage RDF (Lassila et al., 1999) permet d'encoder, d'échanger et de réutiliser des méta-données structurées. Il a été créé pour gérer les méta-données de documents XML mais peut également être utilisé pour des ontologies. Il permet de définir des ressources avec des propriétés et des états. RDF-Schéma définit les relations entre ces ressources. Le pouvoir sémantique de ces deux langages est limité car les axiomes ne peuvent pas être directement décrits. Le type des relations (symétrique, transitive, ...) ne peut être spécifié.

OIL (Ontology Inference Layer) est à la fois un langage de représentation et d'échange pour les ontologies (Decker et al., 2000). Il combine les primitives des langages reposant sur les frames avec une sémantique formelle et des possibilités de raisonnement issues de la logique de description.

DAML+OIL a été proposé par le W3C pour représenter des méta-données et des ontologies (Horrocks et al., 2001). DAML a été transformé en DAML+OIL en intégrant certaines propriétés de OIL. Il repose sur RDF et RDF schéma et fournit en plus des primitives plus riches issues de la logique de description.

OWL (Ontology Web Language) est le standard actuellement proposé par le W3C pour représenter les ontologies. OWL se veut plus représentatif du contenu du Web que XML, RDF et RDF-Schéma en apportant un nouveau vocabulaire avec une sémantique formelle. OWL ajoute du vocabulaire pour décrire les propriétés et classes, comme par exemple la disjonction de classe, la cardinalité (exactement un), l'égalité, les types de propriétés plus riches, les caractéristiques de propriété (symétrie, transitivité, ...) et les classes énumérées.

2. Ontologies existantes modélisant l'image de télédétection

Au cours de ces dernières années, plusieurs ontologies ont été proposées, couvrant différents domaines d'application. Nous dressons un état de l'art sur les ontologies liées au domaine de la géographie et de la télédétection.

2.1. Etat de l'art

2.1.1. L'ontologie INSEE²

C'est une ontologie proposée par l'Institut National de la Statistique et des Etudes Economiques. Elle décrit les événements liés à la création ou à la modification de territoires. En plus, elle est associée à des données issues du code officiel géographique. Ces données concernent les régions, les départements, les arrondissements, les cantons et les communes.

L'ontologie proposée comporte 30 classes avec une profondeur maximale de 5 niveaux, ainsi que 18 propriétés pour les classes, ces propriétés servent aussi d'interrelations pour les données issues du COG. La figure 2.2 présente un extrait de l'ontologie.

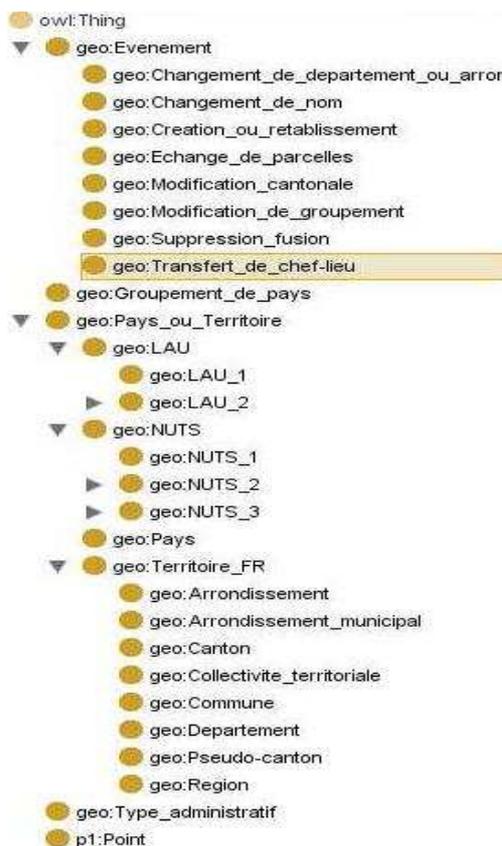


Figure 2.2 : Extrait de l'ontologie INSEE

² <http://rdf.insee.fr/geo/note-uri.html>

2.1.2. L'ontologie Towntology³

Le projet Towntology est issu de la collaboration du LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information) et de l'EDU (Equipe Développement Urbain). Il a pour but de développer une ontologie en langue française dans le domaine de l'aménagement et de l'urbanisme (Roussey et al., 2004).

L'objectif de cette ontologie est de présenter un modèle de représentation des bases de données géographiques, ainsi que la communication entre les différents intervenants dans l'aménagement du territoire. Elle comporte plus de 800 concepts et environ 21 relations sémantiques. Un extrait de l'ontologie est illustré par la figure 2.3.

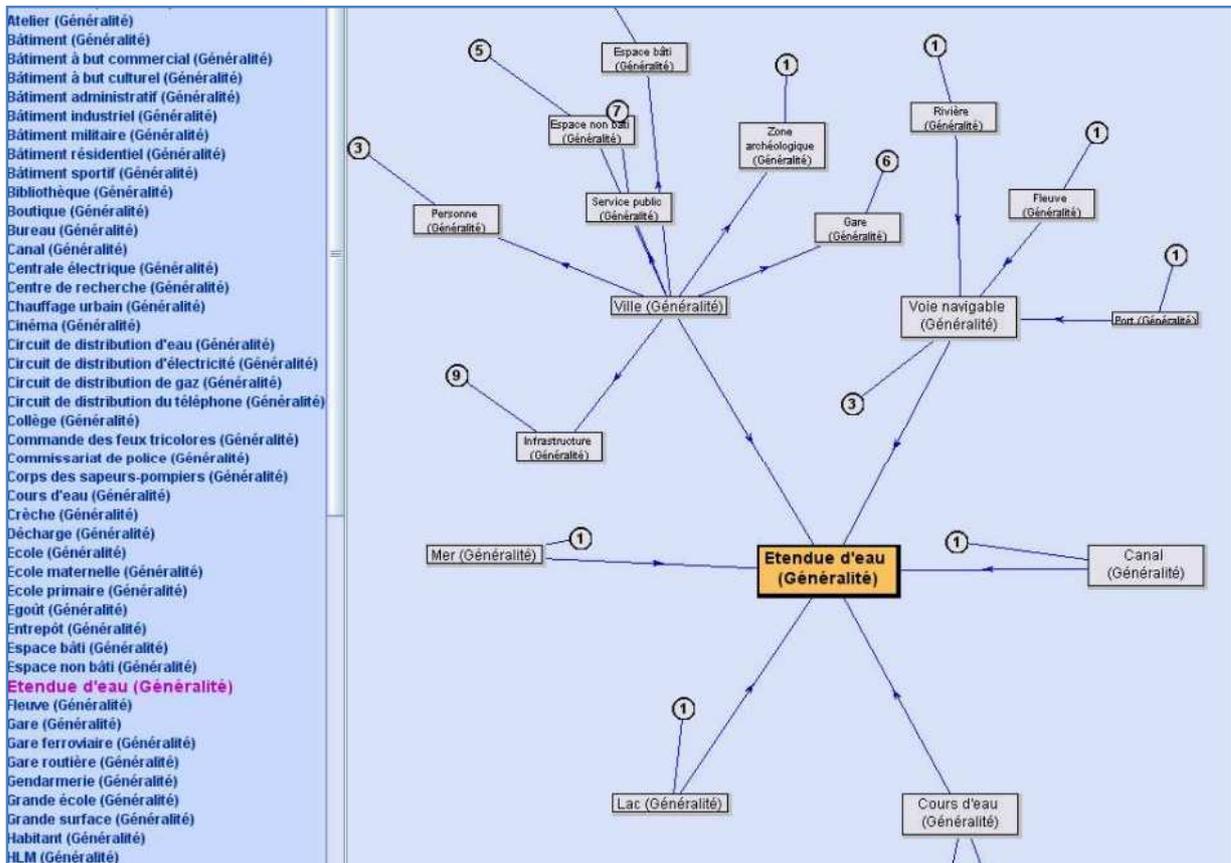


Figure 2.3 : Extrait de l'ontologie de projet towntology

³ <http://liris.cnrs.fr/publis/?id=1594>

2.1.3. L'ontologie de Fodomust⁴

Fodomust est « un projet de fouille de données multi-stratégie pour extraire et qualifier la végétation urbaine à partir de bases de données d'images ». L'ontologie, dont un extrait illustré dans la figure 2.4, a été construite à partir d'entretiens avec les experts du LIV, par apprentissage automatique, et d'un dictionnaire de données recensant les objets géographiques en suivant une méthodologie de conception constituée de 3 phases (spécification, conceptualisation et formalisation). Elle est construite dans le but de pouvoir interpréter et donner du sens aux éléments identifiés sur les images satellitaires (Durant et al., 2007).

```
<owl:Ontology rdf:about="">
  <owl:imports
    rdf:resource="http://protege.stanford.edu/plugins/owl/protege" />
</owl:Ontology>
<owl:Class rdf:ID="Alignement_d_arbres">
  <rdfs:label>Alignement d arbres</rdfs:label>
  <protege:abstract>true</protege:abstract>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Surfaces_arborees" />
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Naturels">
  <protege:abstract>true</protege:abstract>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Plans_d_eau" />
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Batiment">
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Autres_espaces_urbains_specialises" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Espaces_associes_au_reseau_routier" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Decharges_et_depots_de_remblais_
      _deblais" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Zones_industrielles" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Fortes_densite_de_construit_
      _Lotissement" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Emprises_militaires" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Friches_industrielles" />
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Carrieres_actives" />
  </rdfs:subClassOf>
</owl:Class>
```

Figure 2.4 : Extrait de l'ontologie de Fodomust

⁴ <http://lsiit.u-strasbg.fr/afd/sites/fodomust/fr-accueil.php>

2.1.4. L'ontologie Ontogeo-V2-47⁵

C'est une ontologie légère de domaine en langue anglaise s'articulant autour de 1515 concepts qui couvrent (1) les types d'objets géographiques pouvant exister dans une image satellitaire ou dans une carte, et (2) les événements qui peuvent se produire et modifier l'état de ces objets.

Les principaux objets géographiques représentés dans cette ontologie sont : les zones administratives, les constructions, les objets naturels (reliefs, terrains, mers, segments d'eau...), les zones inflammables, les zones d'inondation, les structures géologiques, les zones dangereuses, les substances dangereuses, les réacteurs nucléaires, les objets vulnérables, les voies d'eau (Figure 2.5).

En outre, Ontogeo modélise un ensemble d'événements que peuvent subir ces objets, et décrit les données qui leur sont rattachées, à savoir : des métadonnées, des données spatio-temporelles, des données de localisation et des données thématiques.

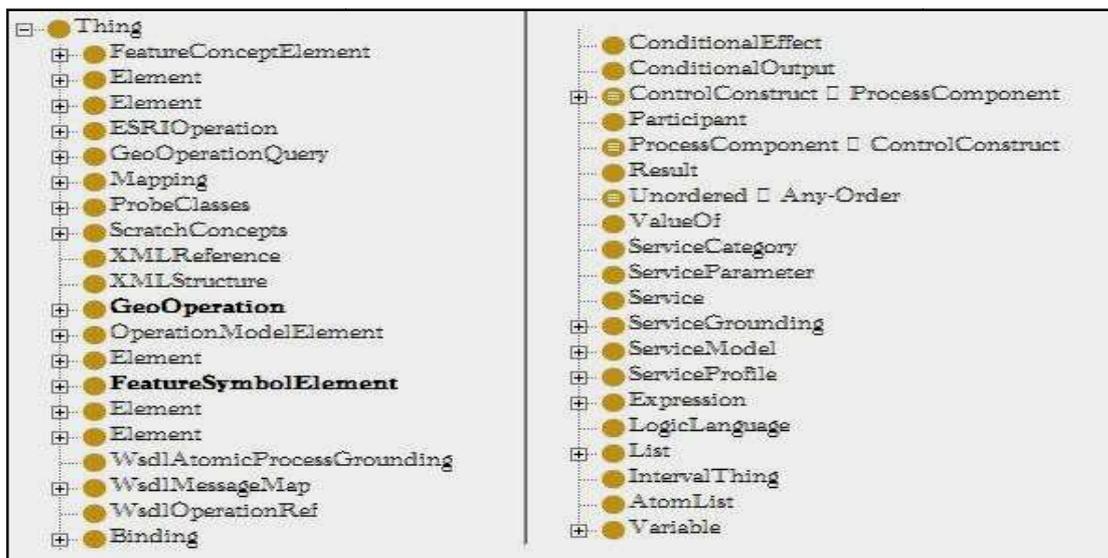


Figure 2.5 : Extrait de l'ontologie d'objets géographiques OntoGeo-V2-47

⁵ <http://geoserver.itc.nl/lemmens/owl/ontogeo-v2-47.owl>

2.1.5. L'ontologie TOP10nl⁶

C'est une ontologie de domaine qui compte 559 concepts, 9 propriétés d'objets et 14 propriétés de données (Figure 2.6). Elle modélise les connaissances relatives au domaine de la topographie et elle s'articule autour de deux grandes familles de concepts : un concept Attribute et ses dérivés et un concept GeoObject et ses dérivés.

- Le concept GeoObject: Il représente les différents types d'objets géographiques. les principaux concepts fils sont : AdministrativeArea, Building, FonctionnelArea, GeographicArea, RailRoadSegment, Relief, RoadSegment, SpecificTerrainElement Terrain, WaterSegment.
- Le concept Attribute : Il décrit les différentes propriétés relatives aux éléments géographiques décrit par le concept GeoObject tels que la surface et le type des zones administratives, le type et la catégorie d'hauteur pour les constructions...



Figure 2.6 : Extrait de l'ontologie d'objets géographiques Top10nl

La figure 2.7 présente les différents types de propriétés de données du concept RoadSegment.

⁶ http://geoserver.itc.nl/lemmens/owl/top10nl_dutch_v1_test.owl

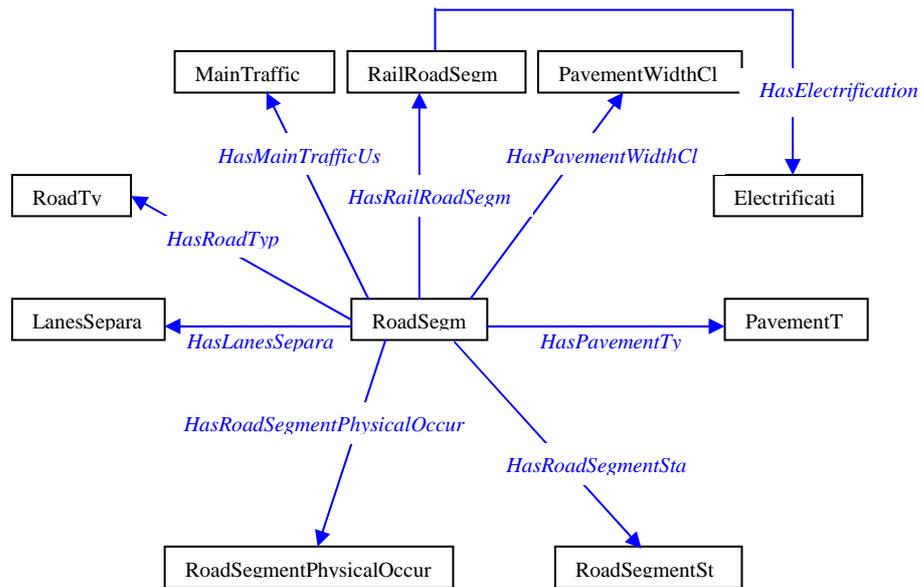


Figure 2.7 : Types de propriétés de données du concept *RoadSegm*

2.1.6. L'ontologie Theme⁷

C'est une ontologie anglaise de domaine en OWL⁸ contenant 141 concepts liés à l'érosion des côtes tels que la pluie, les fleuves, le vent, le niveau de la position de la mer, les inondations, l'infrastructure, les caractéristiques océaniques et terrestres...

Cette ontologie détaille les concepts liés aux éléments suivants : chute de pluie, côtes, leur stabilité et leur type, fleuves et terrains d'eau stagnante ou de sels, types de sol, niveaux de position de mer, topographie. D'un autre côté, l'ontologie propose une panoplie de concepts permettant de modéliser les risques de l'érosion ainsi que les effets des changements côtiers tels que le changement d'habitat et l'augmentation ou l'apparition de plages. Elle renferme aussi des connaissances sur :

- L'environnement physique : les côtes administratives
- Le climat: pluie, vent, stations climatologiques, etc.
- La géologie: la géomorphologie des côtes, la stabilité des côtes, les roches, le sol, et les surfaces géologiques, etc.
- L'hydrologie: les rivières, les inondations, etc.

⁷ <http://mida.ucc.ie/ont/20080124/theme.owl>

⁸ *Web Ontology Language* (OWL) est un langage de représentation des connaissances construit sur le modèle de données de RDF

- Les images satellitaires (LANDSAT, MODIS, SPOT, StableLights) ou aériennes
- L'infrastructure: accès, protection de côtes, transport terrestre, transport marine
- Le paysage
- L'océanographie : caractéristiques océanographiques, niveau de positions des mers, les vagues et marées.
- La topographie.

La figure 2.8 présente un extrait de l'ontologie Theme.



Figure 2.8 : Extrait de l'ontologie Theme

2.1.7. L'ontologie Hydrology⁹

Cette ontologie s'intéresse au domaine de l'hydrographie. C'est une ontologie lourde en langue anglaise écrite dans le langage OWL. Elle a une profondeur de quatre niveaux, elle contient 194 concepts, 16 instances et 43 relations. Un extrait de l'ontologie est présenté par la figure 2.9.

Cette ontologie fait appel à d'autres ontologies : LanguageRelations.owl, SpatialRelations.owl, PoliticalGeography.owl, etc. et elle décrit les caractéristiques topographiques impliquées dans la retenue et le transport d'eau intérieure superficielle. L'ontologie contient des concepts couvrant les aspects suivants : les pentes, les roches, les routes, le bois de construction, les côtes, les bords, les inondations, les plaines inondables, les terrains, les collines, les étendues d'eau artificielle ou naturelle, les sels minéraux, les mouvements de la mer, les moulins à eau, les flux irréguliers, les traces, les pistes, les cours d'eau, les écoulements...

⁹ <http://www.ordnancesurvey.co.uk/oswebsite/ontology/Hydrology/v2.0/Hydrology.owl>



Figure 2.9 : Extrait de l'ontologie Hydrology

2.1.8. L'ontologie de DAFOE¹⁰

Il s'agit d'une ontologie de tâche en RDF contenant 143 concepts et couvrant les données relatives aux régions, aux zones agricoles, à l'occupation des sols, aux forêts, aux constructions diverses, à la végétation, aux zones d'eau, etc. (Charlet et al., 2009). En outre, elle est riche en relations de type inclure, appartenir à, et en relations spatiales qui se répartissent à leur tour en trois grandes classes: de direction, de distance et topologique. La figure 2.10 présente un extrait de l'ontologie.

¹⁰ www.irit.fr/TIA09/thekey/demos/demoDAFOE_TIA.pdf



Figure 2.10 : Extrait de l'ontologie d'objets géographiques DAFOE

2.1.9. L'ontologie FusionTopoCarto2¹¹

C'est une ontologie de domaine, en français, qui compte 761 concepts issus des spécifications des bases de données BDTopo et BDCarto. Les concepts de cette ontologie sont repartis en deux grandes familles : les entités topographiques artificielles et les entités géographiques naturelles. L'importance de cette ontologie réside dans le fait qu'elle est riche en concepts, mais aussi parce qu'elle modélise les entités hydrographiques, topographiques et agricoles (Figure 2.11).

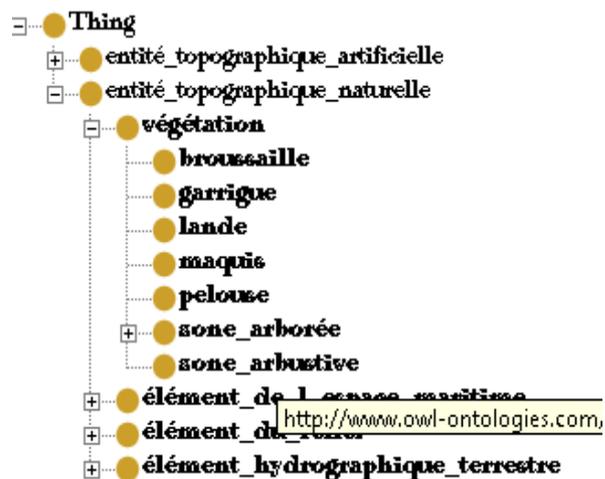


Figure 2.11 : Extrait de l'ontologie d'objets géographiques FusionTopoCarto2

¹¹ fusionTopoCarto2(ontologies-du-cogit).owl

2.1.10. L'ontologie FTT-01¹²

Elle modélise des concepts géographiques qui couvrent les régions, les secteurs administratifs, les colis de terre, les caractéristiques hydrographiques et physio-géographiques. Cependant, cette ontologie se limite à la représentation des relations hiérarchiques entre les concepts et ne définit aucun type de relation spatiale pouvant exister entre les objets constituant une scène donnée. En outre, l'ontologie est dépourvue des propriétés reliant les instances entre elles ou à des données de description. La figure 2.12 présente un extrait de l'ontologie.



Figure 2.12 : Extrait de l'ontologie d'objets géographiques FTT

2.2. Synthèse

Nous présentons ci-dessous un tableau synthétisant les ontologies géographiques.

¹² <http://ifgi.uni-muenster.de/~janowicz/downloads/FTT-v01.owl>.

TOWNTOLGY		EDU (Urban Development Team)		Domaine: l'aménagement et l'urbanisme	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description	
OWL	Ontologie de domaine	Français	- 800 concepts, 21 relations. Profondeur: 3	- Les relations utilisées dans cette ontologie sont de type: est composé de, est un, est situé sur et a pour but.	
DAFOE		DAFOE		Domaine : Objets géographiques	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description	
RDF	Ontologie de tâche	Anglais	- 143 concepts, 64 propriétés, 34 relations. Profondeur: 7	- Données CLC + autres concepts et relations sémantiques : régions, zones agricoles, occupation des sols, forêts, zones en eau, etc. - Relations spatiales : direction, distance et topologie.	
FTT-V01		Université de Münster		Domaine : Géographie	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description	
OWL	Ontologie de domaine	Anglais	- 1262 concepts. - Profondeur: 6	- Concepts géographiques : régions, secteur administratif, caractéristiques hydrographiques et physio géographiques. - Limitation aux relations hiérarchiques.	
HYDROLOGY		Ordnance Survey (OS)		Domaine: Hydrographie	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description	
OWL	Ontologie de domaine	Anglais	- 194 concepts, 16 instances, 43 relations. - Profondeur: 4	- Caractéristiques topographiques impliquées dans la retenue et le transport d'eau intérieure et superficielle : pentes, roches, routes, bois de construction, côtes, bords, plaines inondables, terrains, collines, étendues d'eau artificielle ou naturelle, mouvements de la mer, pistes, cours d'eau... Couvertures des caractéristiques hydrographiques et topographiques	
ONTOGEO		International Institute for Geo-information Science and Earth Observation		Domaine: Objets géographiques	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description	
OWL	Ontologie de domaine	Anglais	- 1515 concepts, 35 propriétés, 193 relations, 265 instances. - Profondeur : 8	Concepts couvrant les types d'objets géographiques pouvant exister dans une image satellitaire ou dans une carte, ainsi que les événements qui peuvent se produire et modifier l'état de ces objets.	
THEME		MIDA - the marine Irish digital atlas http://mida.ucc.ie/		Domaine : Erosions des cotes	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description	
OWL	Ontologie de domaine	Anglais	- 141 concepts, pas de propriétés, instances et relations.	Concepts liés à l'érosion des côtes tels que, la pluie, les fleuves, le vent, le niveau de la position de la mer, les inondations, l'infrastructure, les caractéristiques océaniques et terrestres...	

			Profondeur : 7	
TOP10NL	Lemmens		Domaine : Topographie	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description
OWL	Ontologie de domaine	Multilingue	- 559 concepts, 14 propriétés, 9 relations, et d'instances. Profondeur : 6	- Deux familles de concepts: - Objets géographiques. - Propriétés des objets décrits par les concepts GeoObject et ses fils.
FusionTopoCarto2	Laboratoire COGIT de l'IGN		Domaine : Objets géographiques	Complexité : Légère
Format	Type	Langue	Caractéristiques	Description
OWL	Ontologie de domaine	Multilingue	- 761 concepts, pas de propriétés, de relations et d'instances. Profondeur : 7	- Concepts d'entités géographiques issus des spécifications des bases de données BDTopo et BDCarto. - Deux grandes familles de concepts : Entités topographiques artificielles et Entités géographiques naturelles.

Tableau 2.1 : Synthèse sur les ontologies géographiques

2.3. Discussion

En analysant l'état de l'art, nous faisons ressortir les remarques suivantes :

- 1) Concernant le niveau modélisation, l'information issue de l'image de télédétection est en forte relation avec le capteur d'acquisition et la dimension temporelle.

Toutefois, les ontologies proposées dans la littérature modélisent la dimension spatiale et spectrale (d'un seul capteur), mais pas la *dimension temporelle* ou le *capteur d'acquisition*.

- 2) L'information issue de l'image de télédétection est entachée d'imperfections liées à des sources diverses (le capteur d'acquisition, ...).

Toutefois, bien qu'il existe des travaux qui proposent une modélisation floue des relations spatiales entre les objets d'une image, les imperfections ne sont pas modélisées au niveau des intensions de concepts qui représentent le support de la connaissance.

- 3) La modélisation de la scène nécessite des informations géographiques et des informations issues de l'image (information spectrale, spatiale, et radiométrique). Cependant, l'image satellite n'est pas exploitée en tant que source d'information non négligeable sur la scène, contribuant à alimenter et à enrichir les ontologies.

- 4) Nous remarquons que la majorité des ontologies proposées sont des ontologies de domaine géographique que l'on peut exploiter et réutiliser comme suit :

- L'ontologie FTT modélise un très grand nombre de concepts liés à certains aspects liés au domaine de l'imagerie satellitaire, à savoir : les colis de terre, les caractéristiques hydrographiques et physio géographiques.
- L'ontologie de Fodomust peut servir comme un support consistant pour l'enrichissement de l'ontologie noyau par de nouvelles propriétés extraites.

- La couverture des caractéristiques hydrographiques et topographiques par l'ontologie Hydrology, constitue un complément intéressant des ontologies citées ci-dessus.
 - L'ontologie Hydrology contient des concepts qui aident à l'analyse des risques d'érosion et d'inondation
 - L'ontologie Ontogeo peut servir de base pour l'extraction de certains concepts pouvant être affectés par les événements naturels.
- 5) En plus de la représentation des connaissances d'un domaine, l'ontologie doit offrir des mécanismes de raisonnement inférentiel via les axiomes et les règles qu'elle doit avoir. La majorité des ontologies existantes ne sont pas riches en axiomes et en règles modélisant les connaissances a priori du domaine. D'où la nécessité de formaliser les connaissances du domaine (règles et contraintes).

Nous proposons, dans cette thèse, une ontologie permettant la représentation des connaissances sur les objets de l'image de télédétection, ses caractéristiques visuelles, les relations spatio-temporelles et sémantiques entre les objets ainsi que le mode d'acquisition de l'image, afin de fournir une bonne interprétation au niveau sémantique.

Conclusion

L'ontologie est largement utilisée pour structurer et représenté les connaissances d'un domaine, et notamment le domaine géographique et celui de télédétection. Nous avons ressorti à partir des travaux antérieurs que les ontologies existantes sont plus ou moins complètes pour représenter les connaissances de l'image de télédétection.

Partie B

**Contributions à
l'interprétation d'images de
télédétection par ontologie**

Ce chapitre présente notre contribution concernant la proposition d'une ontologie répondant aux spécificités des images de télédétection. La première section présente le processus de développement de l'ontologie proposée en se basant sur (1) l'enrichissement d'une ontologie de noyau, et (2) la réutilisation des ressources ontologiques existantes. Nous détaillons par la suite les phases du processus de développement. Enfin, nous présentons l'ontologie résultante ainsi que sa définition formelle.

1. Proposition d'une ontologie spatio-temporelle, floue, lourde et évolutive

Afin de remédier aux problèmes rencontrés dans la littérature (cités dans la section 3 du chapitre 2), nous proposons une ontologie ayant une représentation :

- **Spatio-spectro-temporelle** : *modélisant la spatialité, la temporalité et surtout la caractéristique spectrale des images de télédétection. Cette caractéristique est très importante pour une interprétation sémantique d'une image de télédétection.*
- **Floue** : *afin de traiter les imperfections entachant l'image de télédétection. La définition floue concerne la définition des fonctions d'appartenance représentant les valeurs a priori des concepts de l'ontologie.*
- **Lourde** : *afin de doter l'ontologie des mécanismes de raisonnement inférentiel sur un domaine d'application. Cela permet de contribuer à créer des systèmes d'aides à la décision stratégique, à savoir la détection zones à risques d'érosion d'une scène.*
- **Evolutive** : *pour qu'elle soit adaptable et applicable dans plusieurs domaines. Cela permet de réutiliser et exploiter les ressources ontologiques existantes.*

2. Processus de développement de l'ontologie proposée

Afin d'aboutir à notre ontologie, nous proposons un processus de développement basé sur (1) la réutilisation des ressources ontologiques existantes, (2) l'utilisation des images de télédétection comme source d'information, et (3) la formalisation des connaissances *a priori* de domaine (risques naturels, connaissances sur les scènes et capteurs d'acquisition...).

Nous proposons alors de choisir une ontologie, représentant les connaissances de base du domaine, puis de l'adapter au domaine de la télédétection, en l'enrichissant au niveau conceptuel, intensionnel, relationnel et axiomatique, comme illustré dans la figure 3.1. Nous détaillons par la suite chaque étape du processus de développement proposé.

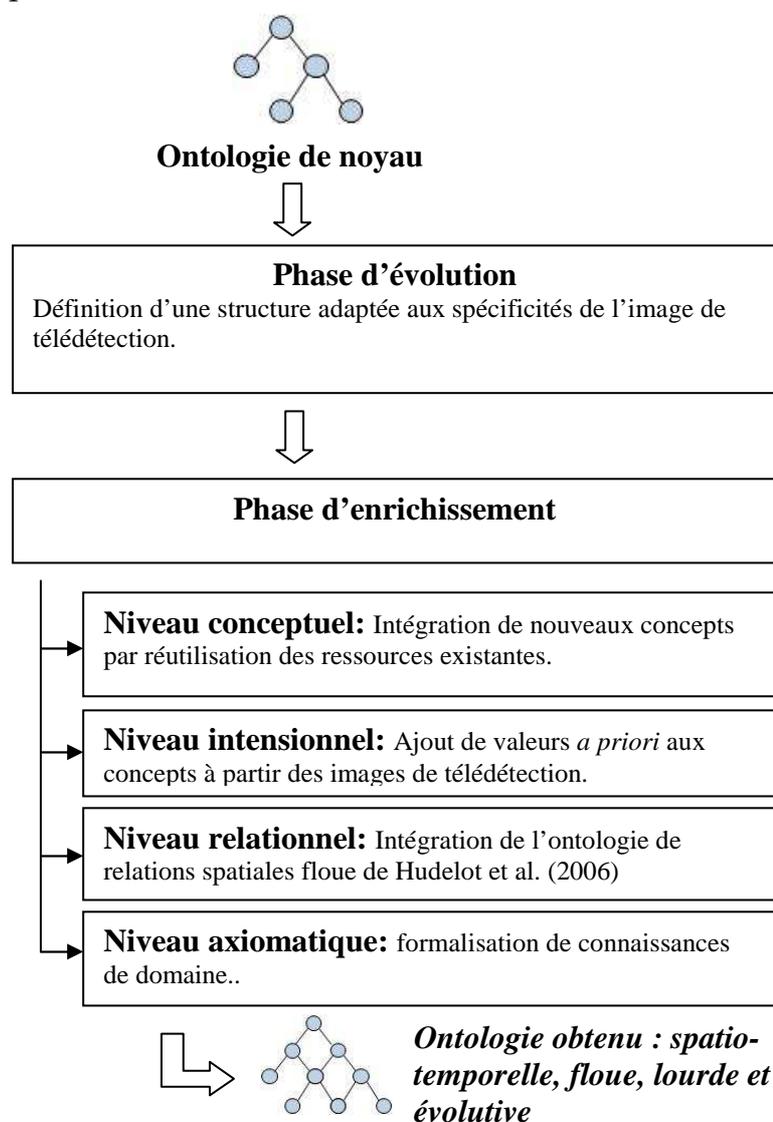


Figure 3.1 : *Processus de développement d'ontologie*

2.1. Choix de l'ontologie de noyau

Le choix de l'ontologie de noyau est très important puisqu'il constitue le point de départ et la base du processus d'enrichissement. L'ontologie de noyau doit modéliser les concepts de base de domaine de télédétection.

D'après l'état de l'art mené sur les ontologies géographiques existantes, nous proposons de réutiliser l'ontologie présentée dans (Durand et al. 2007). Cette dernière se caractérise par sa capacité à modéliser les spécificités spatiales et spectrales des objets détectés dans une scène satellitale, ainsi que sa couverture des différentes orientations liées aux concepts généraux d'une scène urbaine. Pour cela, nous la considérons comme ontologie de noyau. La figure 3.2 présente un extrait de cette ontologie montrant l'hierarchie de concepts et la structure interne des attributs de concepts.

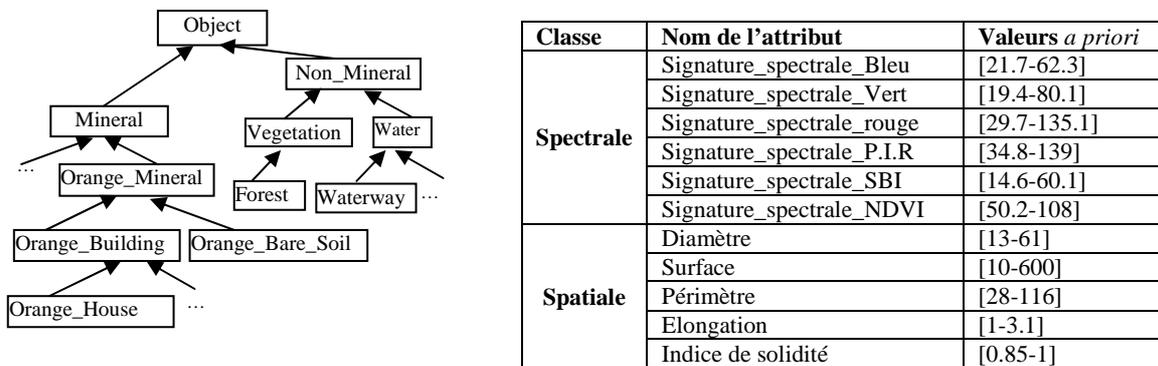


Figure 3.2 : Extrait de l'ontologie de Durand et al. (2007)

2.2. Phase d'évolution de l'ontologie : vers une représentation spectro-spatio-temporelle et floue

Dans l'ontologie de Durand et al. (2007), les concepts sont considérés comme des classes décrivant les régions d'une image. Ils sont définis par un ensemble d'attributs spectraux et spatiaux. En effet, cette structure intensionnelle des concepts est très adaptée aux caractéristiques des images de télédétection.

Cependant, les valeurs radiométriques des régions de l'image dépendent fortement du facteur temporel. Par exemple, un champ de blé peut avoir une couleur verte au mois de janvier, mais jaune au mois d'août. De même, ces valeurs sont définies en fonction du capteur d'acquisition qui se caractérise par sa résolution spectrale. En effet, pour un champ de blé, l'information radiométrique d'une image

SPOT (qui possède trois bandes spectrales) est différente d'une image LandSat possédant sept bandes spectrales. Pour cela, nous proposons de définir les valeurs *a priori* des attributs de chaque concept de l'ontologie en fonction de temps et des bandes spectrales du capteur d'acquisition.

De plus, ces valeurs sont définies par intervalle. Cependant, cette représentation présente un problème d'imprécisions dans la définition des bornes inférieures et supérieures des intervalles, et engendre par la suite des erreurs dans la classification. Illustrons ceci par une région d'une image ayant une valeur radiométrique de « 20.5 », alors, elle ne peut pas être classée à aucune bande pourtant que cette valeur soit très proche de « 20.7 » dans la bande bleu (figure 3.2).

Une solution à ce problème consiste à représenter ces valeurs par les sous-ensembles flous en définissant des fonctions d'appartenance aux attributs de concepts. Le problème qui se pose ici est comment peut-on définir et déterminer ces fonctions d'appartenance.

Dans ce contexte, Ji (2001) a modélisé les classes d'une image de télédétection par des fonctions d'appartenance gaussiennes inspirées de (Yen et Langari, 1999) :

$$\mu_A(x) = e^{-((x-m_A)/\sigma_A)^2} \quad (\sigma_A > 0) \quad (4.1)$$

Avec $\mu_A(x)$ est la fonction d'appartenance à la classe A , m est la moyenne spectrale de la région dans une bande, et σ est la variance. Cette représentation nous paraît pertinente puisqu'elle permet de modéliser l'imprécision entachant les concepts de l'ontologie et de dépasser le problème de représentation par intervalle. Nous proposons alors de l'adopter dans notre contexte, tout en ajoutant la dimension temporelle et celle spectrale. L'équation (4.2) désigne le degré d'appartenance à un attribut α d'un concept c pour une bande spectrale b d'un capteur s à une date t .

$$\mu_{\alpha,s,b,t}(x) = e^{-((x-m_{\alpha,s,b,t})/\sigma_{\alpha,s,b,t})^2} \quad (\sigma_{\alpha,s,b,t} > 0) \quad (4.2)$$

Graphiquement, cette fonction est illustrée comme suit :

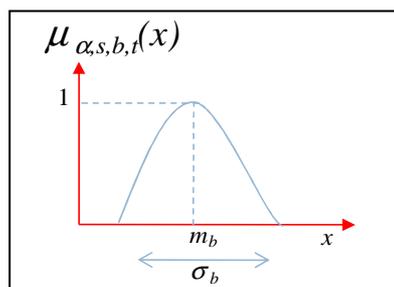


Figure 3.3 : Fonction d'appartenance gaussienne d'un attribut

La figure (3.4) illustre des fonctions d'appartenance gaussiennes modélisant les attributs : radiométrie et INDVI du concept champ de blé, à deux dates : janvier et juillet, et concernant les capteurs : SPOT (bandes XS1 et XS2) et LandSat (bandes TM1 et TM2).

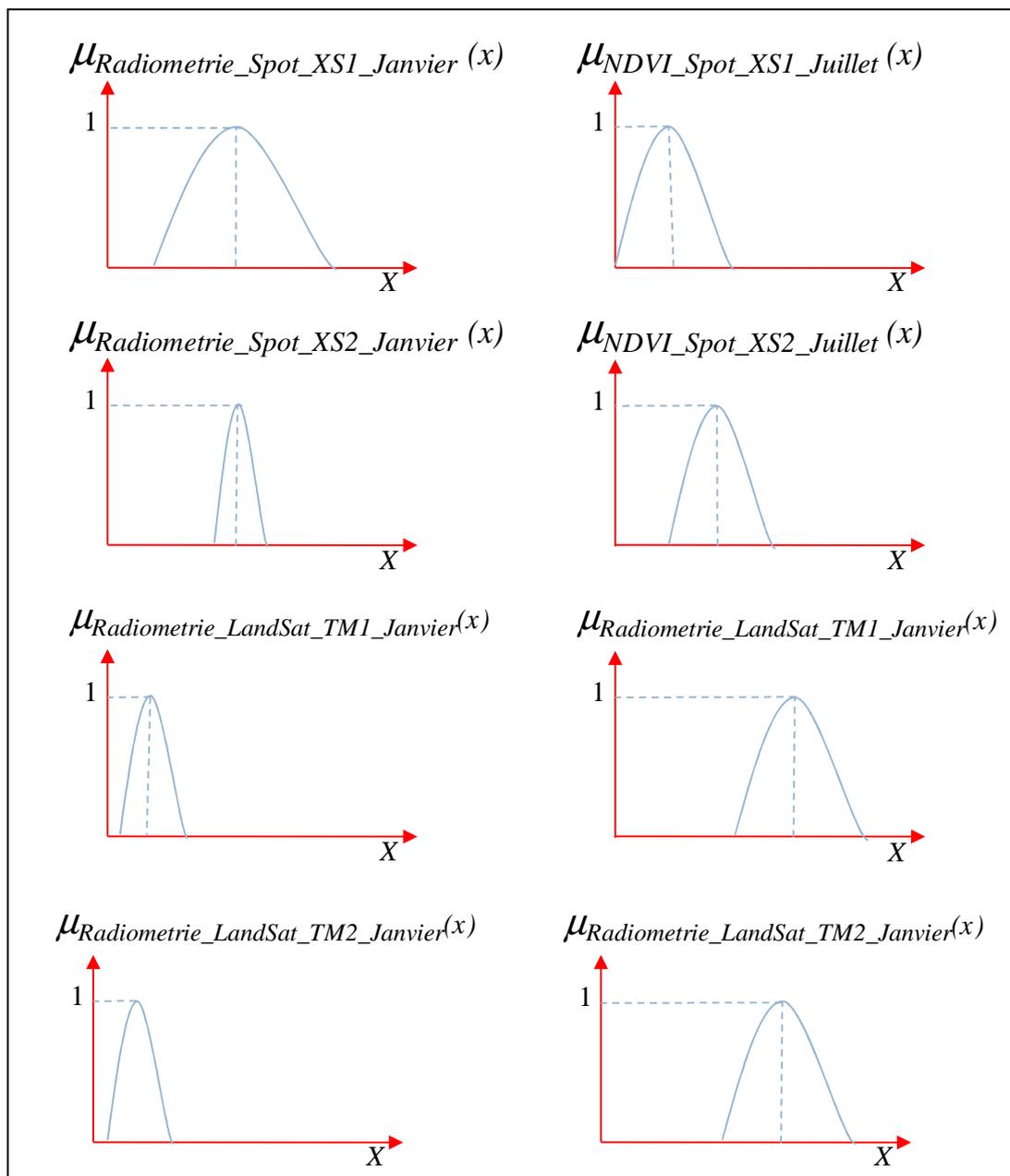


Figure 3.4 : Exemples de fonctions d'appartenance gaussiennes décrivant les attributs d'un champ de blé

Le tableau 4.1 présente la nouvelle structure intensionnelle d'un concept de l'ontologie ayant une représentation *spectro-spatio-temporelle et floue*. Cela permet à l'ontologie de couvrir les spécificités temporelles, spatiales et spectrales des images de télédétection.

Concept « C _i » de l'ontologie							
Attributs	Capteur S ₁			...	Capteur S _n		
	Bande ₁ (b ₁)	...	Bande _n (b _n)		Bande ₁ (b ₁)	...	Bande _n (b _n)
Radiométrie (P ₁)	$t_1: \mu_{P_1_{S_1_{b_1_{t_1}}}}$... $t_n: \mu_{P_1_{S_1_{b_1_{t_2}}}}$...	$t_1: \mu_{P_1_{S_1_{b_n_{t_1}}}}$... $t_n: \mu_{P_1_{S_1_{b_n_{t_2}}}}$		$t_1: \mu_{P_1_{S_n_{b_1_{t_1}}}}$... $t_n: \mu_{P_1_{S_n_{b_1_{t_2}}}}$...	$t_1: \mu_{P_1_{S_n_{b_n_{t_1}}}}$... $t_n: \mu_{P_1_{S_n_{b_n_{t_2}}}}$
NDVI (P ₂)	$t_1: \mu_{P_2_{S_1_{t_1}}}$... $t_n: \mu_{P_2_{S_1_{t_2}}}$				$t_1: \mu_{P_2_{S_n_{t_1}}}$... $t_n: \mu_{P_2_{S_n_{t_2}}}$		
MNDWI (P ₃)	$t_1: \mu_{P_3_{S_1_{t_1}}}$... $t_n: \mu_{P_3_{S_1_{t_2}}}$				$t_1: \mu_{P_3_{S_n_{t_1}}}$... $t_n: \mu_{P_3_{S_n_{t_2}}}$		
SBI (P ₄)	$t_1: \mu_{P_4_{S_1_{t_1}}}$... $t_n: \mu_{P_4_{S_1_{t_2}}}$				$t_1: \mu_{P_4_{S_n_{t_1}}}$... $t_n: \mu_{P_4_{S_n_{t_2}}}$		
Texture (P ₅)	$t_1: \mu_{P_5_{S_1_{b_1_{t_1}}}}$... $t_n: \mu_{P_5_{S_1_{b_1_{t_2}}}}$...	$t_1: \mu_{P_5_{S_1_{b_n_{t_1}}}}$... $t_n: \mu_{P_5_{S_1_{b_n_{t_2}}}}$		$t_1: \mu_{P_5_{S_n_{b_1_{t_1}}}}$... $t_n: \mu_{P_5_{S_n_{b_1_{t_2}}}}$...	$t_1: \mu_{P_5_{S_n_{b_n_{t_1}}}}$... $t_n: \mu_{P_5_{S_n_{b_n_{t_2}}}}$
Surface (P ₆)	$t_1: \mu_{P_6_{S_1_{b_1_{t_1}}}}$... $t_n: \mu_{P_6_{S_1_{b_1_{t_2}}}}$...	$t_1: \mu_{P_6_{S_1_{b_n_{t_1}}}}$... $t_n: \mu_{P_6_{S_1_{b_n_{t_2}}}}$		$t_1: \mu_{P_6_{S_n_{b_1_{t_1}}}}$... $t_n: \mu_{P_6_{S_n_{b_1_{t_2}}}}$...	$t_1: \mu_{P_6_{S_n_{b_n_{t_1}}}}$... $t_n: \mu_{P_6_{S_n_{b_n_{t_2}}}}$
Périmètre (P ₇)	$t_1: \mu_{P_7_{S_1_{b_1_{t_1}}}}$... $t_n: \mu_{P_7_{S_1_{b_1_{t_2}}}}$...	$t_1: \mu_{P_7_{S_1_{b_n_{t_1}}}}$... $t_n: \mu_{P_7_{S_1_{b_n_{t_2}}}}$		$t_1: \mu_{P_7_{S_n_{b_1_{t_1}}}}$... $t_n: \mu_{P_7_{S_n_{b_1_{t_2}}}}$...	$t_1: \mu_{P_7_{S_n_{b_n_{t_1}}}}$... $t_n: \mu_{P_7_{S_n_{b_n_{t_2}}}}$

Tableau 4.1 : Structure proposée d'un concept de l'ontologie

2.3.Phase d'enrichissement de l'ontologie de noyau

Cette phase permet d'enrichir l'ontologie de noyau pour qu'elle couvre bien le domaine de l'imagerie de télédétection. L'enrichissement proposé concerne le niveau conceptuel, intensionnel, relationnel et axiomatique de l'ontologie, et se base sur la réutilisation des ressources ontologiques existantes. Nous aborderons chaque niveau d'enrichissement dans les sous-sections suivantes.

2.3.1. Enrichissement conceptuel par réutilisation de ressources ontologiques

L'enrichissement conceptuel consiste à enrichir l'ontologie de noyau par des nouveaux concepts. Cela implique l'utilisation de méthodes de réutilisation et d'intégration d'ontologies. Nous adoptons alors le processus de réutilisation proposé par Paslaru-Bontas (2005), et composé par les étapes suivantes :

- 1) **Découvertes des ontologies candidates** : identification des ressources ontologiques candidates pour l'enrichissement.
- 2) **Evaluation** : détermination des composants à réutiliser de chaque ressource ontologique.
- 3) **Intégration** : placement des concepts résultants de l'étape précédente tout en vérifiant la cohérence de l'ontologie de noyau.

La figure 3.5 illustre ces étapes qui seront détaillées dans les prochaines sous-sections.

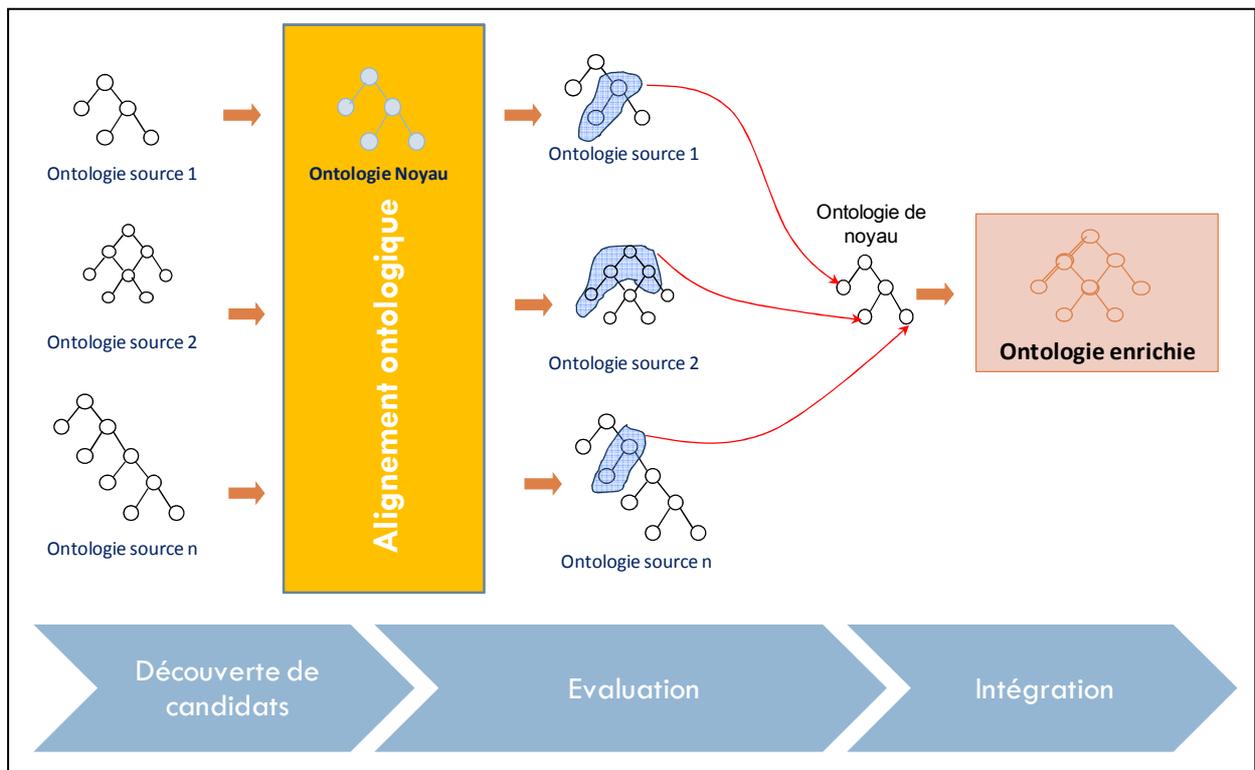


Figure 3.5 : *Processus d'enrichissement conceptuel*

2.3.1.1. Etape 1 : Découverte des ontologies candidates

Les ressources ontologiques existantes servent comme support essentiel pour l'enrichissement de l'ontologie de noyau. Ces ressources sont choisies en fonction de leur richesse en termes de concepts et de leur liaison au domaine de l'image de télédétection. En effet, ces ontologies seront comparées à l'ontologie de noyau par le calcul d'un score de similarité globale. Seules les ontologies vérifiant une similarité supérieure à un seuil fixé par un expert sont retenues pour l'alignement.

2.3.1.2. Etape 2 : Evaluation

Cette étape permet de déterminer le/les concept(s) réutilisable(s) de chaque ontologie retenue de l'étape précédente. Pour cela, nous appliquons un processus d'alignement entre l'ontologie de noyau à enrichir, notée O_{cible} , et chaque ressource candidate, notée O_{source} .

Ce processus consiste à produire un ensemble de correspondances entre les concepts des ontologies sources et ceux de l'ontologie cible. Ces correspondances peuvent être des relations d'équivalence ou de subsumption. Pour ce faire, nous utilisons des techniques d'alignement terminologiques, sémantiques et structurelles. Les résultats d'alignement nécessitent ainsi un filtrage pour pallier aux problèmes de conflit et de redondance affectant les règles de correspondance extraites (Figure 3.6). Nous détaillons par la suite les techniques d'alignement et de filtrage, ainsi que la propriété itérative du processus.

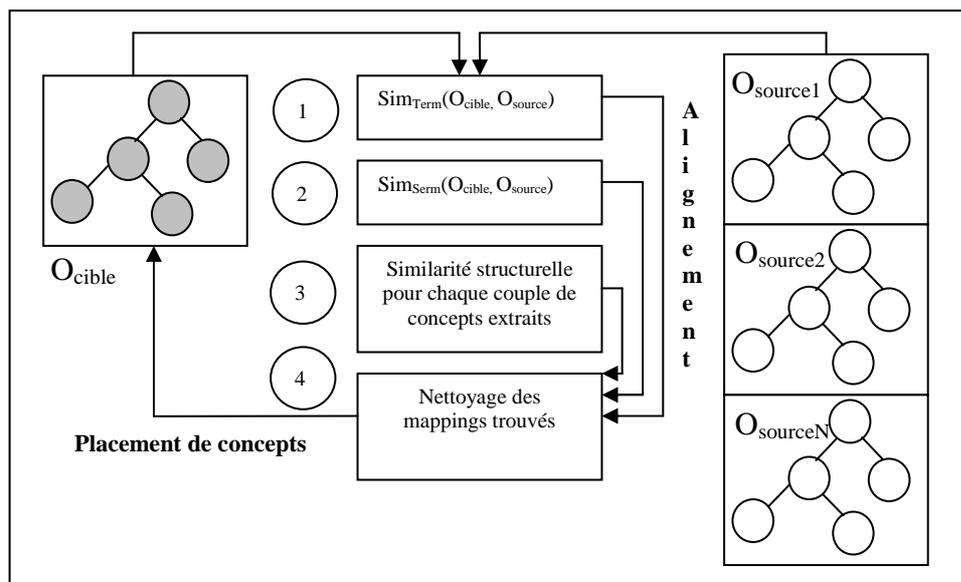


Figure 3.6 : Alignement, filtrage, et placement entre O_{cible} et O_{source}

a) Alignement entre O_{cible} et O_{source}

Nous présentons les techniques d'alignement utilisées pour établir les mappings existantes entre O_{cible} et O_{source} .

Les techniques terminologiques : Elles consistent à effectuer des mesures de similarité sur les identificateurs et les labels caractérisant les couples d'entités à appairer. Les relations pouvant être déduites par l'application de cette technique sont l'équivalence et la subsumption. Afin de s'assurer de la cohérence de ces relations,

une mesure structurelle est calculée pour chacune des correspondances trouvées. Selon cette mesure, une relation est jugée consistante, si ses deux extrémités sont proches structurellement, c'est-à-dire, ont des ressemblances au niveau de leurs voisinages directs.

L'égalité stricte des labels relatifs aux deux concepts comparés implique une relation d'équivalence. Dans le cas de l'inégalité, c'est la relation d'inclusion de labels qui doit être vérifiée. L'existence d'une telle relation implique une relation de subsomption entre les concepts appariés.

Pour ce faire, nous nous basons sur des heuristiques terminologiques simples qui portent sur les labels et les synonymes des termes désignant les concepts.

Edit-Distance : Si les labels de deux concepts ne se différencient pas par plus de deux caractères, alors les concepts considérés peuvent être reliés par une relation d'équivalence.

Inclusion de labels : Si tous les mots de l'identificateur d'un concept A sont inclus dans les mots d'un concept B, alors B sera considéré comme plus spécialisé que A (B isA A).

Les techniques sémantiques : Elles consistent à chercher des relations sémantiques possibles en s'appuyant sur la base lexicale WordNet¹.

WordNet est une ressource lexicale de langue anglaise qui regroupe des termes (noms, verbes, adjectifs et adverbes) en ensembles de synonymes appelés synsets. Un synset regroupe tous les termes dénotant un concept donné. Le terme associé à un concept est représenté sous une forme lexicalisée, sans marque de féminin ni de pluriel. Les synsets sont reliés entre eux par des relations sémantiques : relation de généralisation/spécialisation et relation composant/composé.

Les techniques basées sur les chaînes de caractères ne sont pas suffisantes quand les concepts sont sémantiquement proches et quand leurs noms sont différents, l'interrogation d'une ressource linguistique tel que WordNet peut indiquer que les concepts sont similaires.

¹ <http://wordnet.princeton.edu/>

L'alignement utilisant des ressources complémentaires comme WordNet consiste à rechercher les mappings entre les termes selon l'hypothèse suivante : Deux termes T_X et T_Y sont sémantiquement similaires s'il existe un synset S' de T_X qui soit aussi un synset de T_Y ou qui soit relié à un synset de T_Y par une suite de relations de généralisation/spécialisation dans WordNet. Un mapping T_X isA T_Y sera rejeté si S' n'existe pas ou s'il existe mais n'est pas sémantiquement similaire à l'un des ancêtres de T_X dans son ontologie d'origine.

Il s'agit donc de chercher pour chaque concept de l'ontologie source, les synsets qui correspondent à ses termes associés. S'il existe un synset en commun entre les deux termes désignant le couple de concepts (source, cible), c'est alors une relation d'équivalence qui est générée et ajoutée à la liste des couples (X_{source}, Y_{cible}) équivalents. Sinon, un balayage s'effectuant sur les généralisations et les spécialisations du terme T_X pour voir s'il peut être relié par une relation de subsomption à T_Y . Une fois le parcours est terminé sur tous les concepts des ontologies sources, les correspondances trouvées subissent un travail de filtrage conduisant à l'élimination des conflits et de redondances.

Les techniques structurelles : Afin de s'assurer de la proximité sémantique des concepts appariés et pour pallier les problèmes liés à la polysémie, une vérification au moyen d'une technique structurelle semble intéressante.

Les relations d'équivalence trouvées suite à l'application des techniques terminologiques et sémantiques, serviront pour deux raisons essentielles :

- Nettoyage des mappings de subsomption : les mappings qui apparaissent à la fois dans des relations de subsomption et d'équivalence, sont confrontés à une mesure structurelle qui conduit au rejet de celui ayant la valeur la plus petite.
- Extraction des relations de subsomption : les descendants d'un concept source sont considérés comme étant des fils possibles (candidats) pour son équivalent dans l'ontologie cible. A cet égard, nous proposons d'ajouter, tous les paires de concepts (concept cible, descendant du concept source) à la liste des mappings de subsomption. Ces relations d'équivalence serviront comme une troisième source pour l'extraction des relations de subsomption.

L'utilisation de cette diversité de techniques d'alignement ainsi que l'utilisation de plusieurs ontologies sources aide à fouiller les concepts qui peuvent enrichir l'ontologie noyau. Cependant, cet apport ne cache pas l'existence d'un véritable

problème de conflits et de redondances entre les mappings d'un même type de relations, d'un côté, et entre les mappings de deux relations, d'un autre côté.

Les techniques structurelles permettent de pallier une importante partie de ces problèmes. Pour la mesure de similarité structurelle entre deux concepts, nous proposons d'utiliser les deux techniques suivantes:

- **Taxonomic Overlap [Maedche et Staab 02]** : Cette technique aide à déterminer le degré d'importance d'un même concept dans les deux ontologies à aligner. Elle est basée sur la « *Sémantique Cotopy* » d'un concept C dans une ontologie O :

$SC(c, O)$ = l'ensemble de ses descendants et généralisants

La mesure *Taxonomic Overlap* correspond, selon l'existence ou non d'un même concept dans les deux ontologies à apparier, correspond à l'une des quantités suivantes:

$TO (C_i, H_1, H_2) = \frac{|SC(C_i, H_1) \cap SC(C_i, H_2)|}{|SC(C_i, H_1) \cup SC(C_i, H_2)|}$, dans le cas où le concept C_i existe dans l'ontologie H_2 .

$TO' (C_i, H_1, H_2) = \text{Max}_{C_i \in H_2} \left[\frac{|SC(C_i, H_1) \cap SC(C_i, H_2)|}{|SC(C_i, H_1) \cup SC(C_i, H_2)|} \right]$, dans le cas où le concept C_i n'existe pas dans l'ontologie H_2 .

L'avantage principal de cette technique est qu'elle rapproche deux concepts portant des labels différents bien qu'elle se base sur des comparaisons syntaxiques. Il s'agit en effet de vérifier si le concept en question est entouré (descendants et généralisants) par des concepts similaires dans les deux ontologies à apparier. Elle permet de rapprocher par exemple les deux concepts " Surface d'eau" et " Entité Hydrographique de surface" vu qu'il y a les fils "lac", "marais" et "mare" en commun entre les deux.

- **Mesure de Wu & Palmer** : La similarité entre deux concepts $c1 \in O_1$, et $c2 \in O_2$ dépend de la similarité entre leurs voisinages. En effet, si le voisinage d'un concept $c1$ noté $V(c1)$ est similaire au voisinage du concept $c2$ noté $V(c2)$ alors $c1$ et $c2$ sont similaires d'une certaine manière. Le voisinage d'un concept c d'une ontologie O est défini par :

$V(c) = \{ \text{tout concept } c' \in O / c' = \text{père}(c) \text{ ou } c' \in \text{fils}(c) \};$

La mesure structurelle est déterminée par la formule suivante :

$$Sim_{str}(c1, c2) = \frac{\sum_{(rc1, rc2) \in VR(c1, c2)} Sim_{ling}(rc1, rc2) * Sim_{wp}(c1, c2) * (1 - |d1 - d2|^2)}{|VR(c1, c2)|}$$

- Où : $Simling(rc1,rc2)$: définit la similarité linguistique entre les concepts $rc1$ et $rc2$ qui correspondent aux voisinage relatifs respectifs des concepts $c1$ et $c2$.

$$Sim_{wp}(c1,c2) = \frac{profond(LCA(c1,c2))}{profond(c1) + profond(c2)}$$

- $LCA(c1,c2)$ est le petit ancêtre commun de $c1$ et $c2$.
- $profond(LCA(c1,c2))$ correspond au nombre d'arcs séparant $LCA(c1,c2)$ de la racine.
- $profond(c1)$ désigne la longueur du chemin séparant $c1$ de la racine en passant par $LCA(c1,c2)$.
- $d1 = simwp(c1,rc1)$ et $d2 = simwp(c2,rc2)$
- $VR(c1,c2)$: Voisinage relatif correspondant au voisinage qui possède un appariement avec un voisinage de la même position dans l'ontologie cible (père - père, fils-fils).

b) Filtrage des mappings

Afin de résoudre les problèmes de conflits et de redondances qui peuvent être rencontrés entre les mappings d'un même type de relations ou de deux types de relations différents, nous proposons un ensemble règles permettant un filtrage de ce mappings :

- **Règle 1** : Si la même correspondance apparaît deux fois, une dans les relations d'équivalence, et une dans les relations de subsomption, c'est la relation possédant la similarité structurelle la plus grande qui est retenue dans la liste finale de mappings.
- **Règle 2** : Si un concept source apparaît dans plusieurs relations de subsomption, c'est la relation possédant la plus grande similarité structurelle qui est préservée.
- **Règle 3** : Si les deux extrémités d'une relation de subsomption apparaissent inversement dans deux mappings, et si leurs similarités sont très proches, alors elles seront regroupées dans une seule relation d'équivalence qui sera dotée de la mesure structurelle la plus élevée. Si la différence entre les similarités relatives aux deux correspondances dépasse un seuil donné, alors c'est le sens ayant la proximité structurelle la plus grande qui est retenu.

- **Règle 4** : Un concept source subsumant candidat ne peut être validé que s'il n'a pas d'équivalent dans les descendants de son correspondant relatif à l'ontologie cible. Ce concept ne doit pas non plus apparaître comme un fils pour les différents descendants de son correspondant cible.

c) Qualité itérative du processus d'alignement

Afin de tirer profit au maximum des potentialités du processus d'enrichissement, nous proposons d'utiliser itérativement une multitude d'ontologies sources dans le but d'enrichir au mieux notre ontologie cible. Cela permet en effet de :

- Découvrir des concepts qui n'ont pas pu être découverts dans des itérations antérieures. En effet, l'extension de l'ontologie cible par les concepts d'une ontologie source conduit à l'accroissement de la similarité des voisins entre les concepts cibles enrichis par des nouveaux descendants et leurs correspondants dans l'ontologie source. Par conséquent, la mesure structurelle d'un mapping peut augmenter d'une itération à l'autre.
- Augmenter le niveau de détail pour un domaine particulier d'une itération à une autre. En effet, les relations de subsumption générées à partir des relations d'équivalences, se limitent aux descendants directs des concepts sources. Le fait d'itérer sur cette même ontologie aide à découvrir les subsumants des concepts ajoutés récemment comme des descendants à des concepts cibles dans des itérations antérieures, et par conséquent de descendre dans la hiérarchie des concepts sources.
- L'utilisation d'une variété d'ontologies sources permet à l'ontologie cible de couvrir plusieurs types d'objets liés à l'imagerie satellitaire et à l'étude des risques naturels. Il s'agit, en effet, d'extraire des concepts faisant partie de plusieurs domaines tels que l'hydrographie, la végétation, la topographie, l'urbanisme, etc.

2.3.1.3. Etape 3 : Placement

Cette étape permet d'ajouter les mappings retenues de l'étape précédente (filtrées et validées) dans l'ontologie de noyau. Ces mappings peuvent être des subsumptions ou des équivalences.

Nous présentons par la suite l'algorithme général du processus d'enrichissement conceptuel :

Algorithme proposé d'enrichissement conceptuel :

Entrées

O_{cible} : Ontologie de départ ou noyau.
 $O_{source\ 1}, \dots, O_{source\ N}$: Ontologie candidate pour l'extension de O_{cible} .
 WordNet : base lexicale servant pour les mappings sémantiques.
 Seuil : Seuil de similarité structurelle fixé par l'expert.

Variables intermédiaires

Y_{cible} : Concept de l'ontologie cible O_{cible} .
 X_{source} : Concept de l'ontologie source O_{source} .
 $Liste_{isEquiv}$: Liste des couples (Y_{cible}, X_{source}) reliés par une relation d'équivalence

Résultats

$Liste_{isA}$: Liste des couples (X_{source}, Y_{cible}) tel que Y_{cible} *est un* X_{source}

Fonctions

$Sim_{globale}(O_{source}, O_{cible})$: retourne la similarité structurelle globale entre les deux ontologies O_{cible} et O_{source} .
 $Sim_{lables}(X_{source}, Y_{cible})$: Retourne un booléen qui indique si les lables de Y_{cible} et X_{source} sont identiques ou non.
 $Sim_{Wu\&Palmer}(X_{source}, Y_{cible}, O_{source}, O_{cible})$: calcule la similarité structurelle entre les concepts Y_{cible} et X_{source} .
 $Inclusion_{Labels}(X_{source}, Y_{cible})$: retourne vrai si les labels de Y_{cible} sont inclus dans ceux de X_{source} , faux sinon.

Début

Pour chaque couple d'ontologies (O_{source}, O_{cible}) **Faire**

Si $Sim_{globale}(O_{source}, O_{cible}) \geq \text{seuil}$ **Alors**

Pour tout concept $Y_{cible} \in O_{cible}$ Et $X_{source} \in O_{source}$ **Faire**

Si $Inclusion_{Labels}(X_{source}, Y_{cible})$ et $Sim_{Wu\&Palmer}(X_{source}, Y_{cible}, O_{source}, O_{cible}) \geq \text{seuil}$ **Alors**

Ajouter le couple (X_{source}, Y_{cible}) à la liste de correspondances de subsomption $Liste_{isA}$

Sinon Si $Sim_{lables}(X_{source}, Y_{cible})=1$ Et $Sim_{Wu\&Palmer}(X_{source}, Y_{cible}, O_{source}, O_{cible}) \geq \text{seuil}$ **Alors**

Ajouter le couple (X_{source}, Y_{cible}) à la liste de correspondances de d'équivalence $Liste_{isEquiv}$

Fin Si

Chercher une correspondance sémantique entre X_{source} et Y_{cible} dans Wordnet

S'il existe une relation sémantique de subsomption entre X_{source} et Y_{cible} et similarité structurelle $Sim_{Wu\&Palmer}(X_{source}, Y_{cible}, O_{source}, O_{cible}) \geq \text{seuil}$ **Alors**

Ajouter le couple (X_{source}, Y_{cible}) à la liste de correspondances $Liste_{isA}$

Sinon Si X_{cible} est synonyme de Y_{cible} Et similarité structurelle $Sim_{Wu\&Palmer}(X_{source}, Y_{cible}, O_{source}, O_{cible}) \geq \text{seuil}$ **Alors**

Ajouter le couple (X_{source}, Y_{cible}) à la liste de correspondances de d'équivalence $Liste_{isEquiv}$

Fin Si

Fin Pour

Pour chaque fils de $X_{source} \in Liste_{isEquiv}$ **Faire**

Ajouter les couples (fils de X_{source}, Y_{cible}) à la liste de correspondances $Liste_{isA}$

Fin Pour

Filtrage des mappings trouvés

Placement des concepts sources restants dans O_{cible}

Fin Pour

Retourner O_{cible}

Fin

2.3.2. *Enrichissement intensionnel de concepts à partir d'images*

Cette phase d'enrichissement est très intéressante pour que l'ontologie soit riche en connaissances liées au domaine de l'image de télédétection. Elle consiste à alimenter les intensions des concepts par des valeurs *a priori* représentées par des fonctions d'appartenances gaussiennes, tout en respectant la structure intensionnelle de concepts proposée.

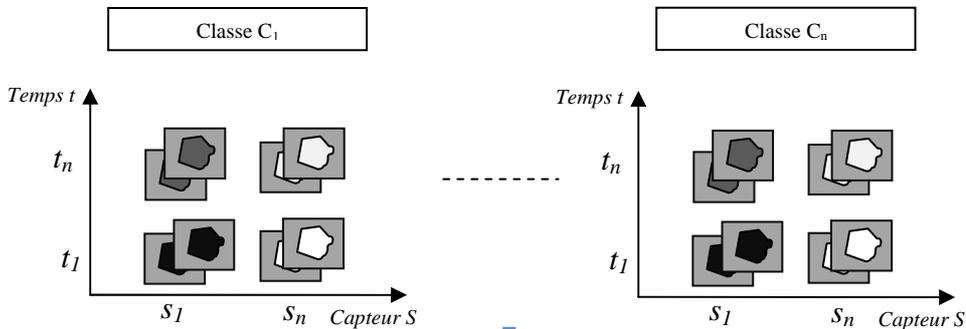
Ce processus d'enrichissement intensionnel est basé sur l'extraction des informations spectrales et visuelles de l'image de télédétection (figure 3.7). Considérons un ensemble d'images multi-dates et multi-capteurs représentant une classe, nous essayons d'extraire les informations spectrales de l'image à savoir la radiométrie et les indices de végétation, d'eau et de brillance, et celles visuelles tel que la texture et la forme. Puis, nous déterminons les paramètres des fonctions d'appartenances gaussiennes pour chaque image à savoir la moyenne m et la variance σ . Enfin, les concepts de l'ontologie de noyau O_{cible} seront enrichis par ces fonctions d'appartenances.

2.3.3. *Enrichissement relationnel*

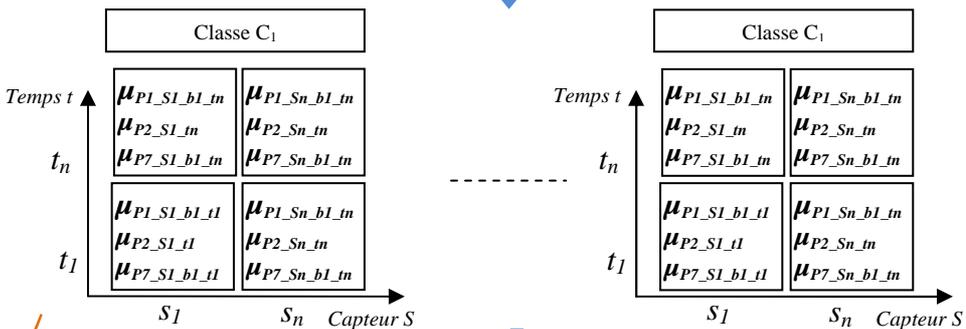
Les relations spatiales entre les objets d'une image jouent un rôle important dans l'interprétation de scènes satellitaires. Freeman (1975) a proposé treize relations spatiales spécifiant la position relative d'une région par rapport à une autre. Ces relations peuvent être classées en trois familles :

- **Les relations topologiques** : à l'intérieur de, à l'extérieur de, touche, entre, devant, à côté de, derrière ...
- **Les relations liées à la distance** : loin de, près de ...
- **Les relations directionnelles** : au-dessus de, en-dessous de, à gauche de, à droite de...

Imagettes multi-dates et multi-capteurs représentant une classe



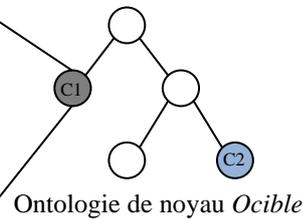
Détermination des fonctions d'appartenances gaussiennes



Enrichissement des intensions de concepts

Enrichissement intensionnel

Attributs	Concept « C ₁ » de l'ontologie			
	Capteur S ₁		Capteur S _n	
	Bande ₁ (b ₁)	... Bande _l (b _l)	Bande ₁ (b ₁)	... Bande _l (b _l)
Radiométrie (P ₁)	$f_{11}^1: \mu_{P_1 S_1 b_1 t_1}$...	$f_{1l}^1: \mu_{P_1 S_1 b_l t_1}$...	$f_{11}^n: \mu_{P_1 S_n b_1 t_1}$...	$f_{1l}^n: \mu_{P_1 S_n b_l t_1}$...
NDVI (P ₂)	$f_{21}^1: \mu_{P_2 S_1 t_1}$...	$f_{2l}^1: \mu_{P_2 S_1 t_l}$...	$f_{21}^n: \mu_{P_2 S_n t_1}$...	$f_{2l}^n: \mu_{P_2 S_n t_l}$...
MNDWI (P ₃)	$f_{31}^1: \mu_{P_3 S_1 t_1}$...	$f_{3l}^1: \mu_{P_3 S_1 t_l}$...	$f_{31}^n: \mu_{P_3 S_n t_1}$...	$f_{3l}^n: \mu_{P_3 S_n t_l}$...
SBI (P ₄)	$f_{41}^1: \mu_{P_4 S_1 t_1}$...	$f_{4l}^1: \mu_{P_4 S_1 t_l}$...	$f_{41}^n: \mu_{P_4 S_n t_1}$...	$f_{4l}^n: \mu_{P_4 S_n t_l}$...
Texture (P ₅)	$f_{51}^1: \mu_{P_5 S_1 b_1 t_1}$...	$f_{5l}^1: \mu_{P_5 S_1 b_l t_1}$...	$f_{51}^n: \mu_{P_5 S_n b_1 t_1}$...	$f_{5l}^n: \mu_{P_5 S_n b_l t_1}$...
Surface (P ₆)	$f_{61}^1: \mu_{P_6 S_1 b_1 t_1}$...	$f_{6l}^1: \mu_{P_6 S_1 b_l t_1}$...	$f_{61}^n: \mu_{P_6 S_n b_1 t_1}$...	$f_{6l}^n: \mu_{P_6 S_n b_l t_1}$...
Périmètre (P ₇)	$f_{71}^1: \mu_{P_7 S_1 b_1 t_1}$...	$f_{7l}^1: \mu_{P_7 S_1 b_l t_1}$...	$f_{71}^n: \mu_{P_7 S_n b_1 t_1}$...	$f_{7l}^n: \mu_{P_7 S_n b_l t_1}$...



Structure intensionnelle du concept C_1

Figure 3.7 : Processus d'enrichissement intensionnel des concepts à partir d'images de télédétection

Plusieurs approches ont été proposées permettant la représentation qualitative et quantitative de ces relations. Citons, à titre d'exemple, l'algèbre RCC (*Region Connection Calculus*), proposé par Clarke (Clarke, 1981), et permettant de représenter qualitativement les relations topologiques (tableau 4.2).

Relation spatiale	Description	Représentation graphique
$DC(x,y)$	x est déconnecté de y	
$EC(x,y)$	x est connexe à y	
$PO(x,y)$	x recouvre partiellement y	
$TPP(x,y)$	x est une partie propre tangentielle de y	
$NTPP(x,y)$	x est une partie propre non tangentielle de y	
$TPPi(x,y)$	x possède pour partie propre tangentielle y	
$NTPPi(x,y)$	x possède pour partie propre non tangentielle y	
$EQ(x,y)$	x est égal à y	

Tableau 4.2 : Relation spatiales topologiques RCC-8

Le formalisme des ensembles flous a été proposé pour représenter les relations spatiales afin de modéliser l'imprécision intrinsèque de relations telles que «proche de» ou encore «très loin». Cette approche consiste à définir des fonctions d'appartenance représentant le degré de satisfaction de la relation par rapport à un objet de référence en tout point de l'espace, nommé *paysage flou* (Bloch, 1995 ; Bloch, 1999). L'estimation de ces fonctions d'appartenance est fondée sur la morphologie mathématique modélisant principalement les relations de direction, l'adjacence flou et les relations de distances (figure 3.8).

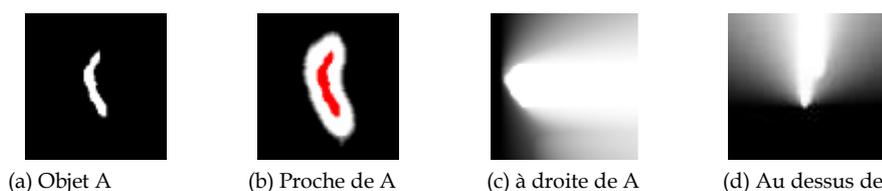


Figure 3.8 : Exemples de «paysage flou» représentant les relations de distances et de direction

Dans ce contexte, Hudelot et al. (2008) ont proposé une ontologie modélisant les relations spatiales (Figure 3.9). Cette ontologie présente une organisation hiérarchique des relations spatiales selon leurs caractéristiques. En outre, elle est enrichie par une représentation floue fondée sur la morphologie mathématique et les logiques de description, et permettant, par la suite, de maximiser le raisonnement spatial en utilisant les opérateurs flous (t-norme, t-conormes, etc.). Comme application, les auteurs ont utilisé l'ontologie pour décrire sémantiquement l'anatomie cérébrale.

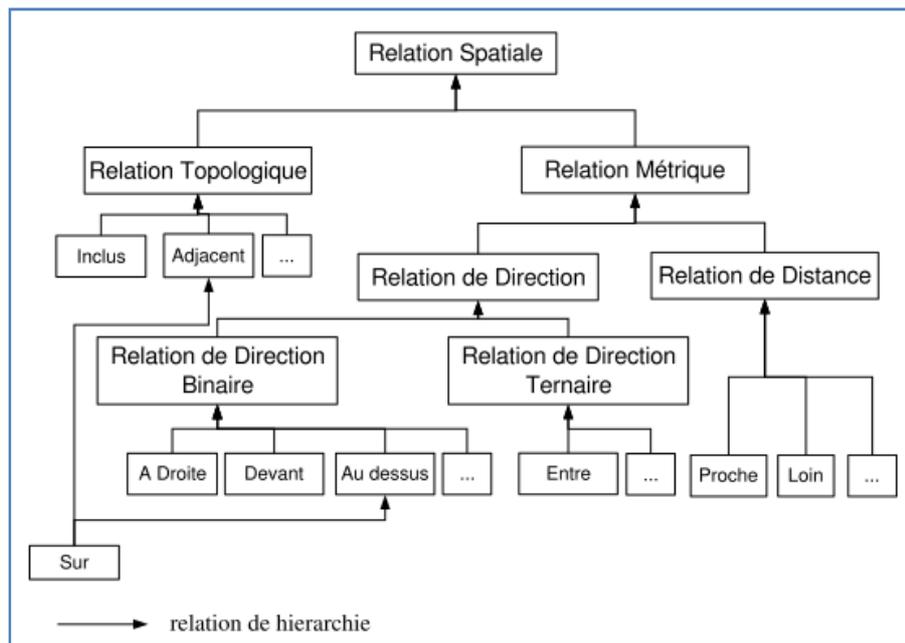


Figure 3.9 : Extrait de l'ontologie de relations spatiales floues (Hudelot et al., 2008)

En effet, cette ontologie nous paraît intéressante et très proche de notre domaine d'application. On propose, alors, de la réutiliser et de l'intégrer dans notre ontologie, cela permet de favoriser l'interprétation sémantique dans les images de télédétection.

2.3.4. Enrichissement axiomatique

L'image satellitale est très dépendante du capteur d'acquisitions : en effet, l'image du capteur radar est totalement différente de celle du capteur optique même si elles représentent la même scène. Nous proposons alors d'ajouter un ensemble d'axiomes modélisant les connaissances sur les capteurs (type, résolution spatiale, nombre de bandes spectrales,...). Nous proposons aussi d'ajouter un ensemble

d'axiomes et de règles modélisant les connaissances sur les phénomènes et les risques naturels en fonction de concepts de l'ontologie et de leurs relations spatiales.

Exemple de connaissances de domaine sur le risque d'érosion :

Si une région est constituée de matériau parental moyennement cohérent (marno-calcaire), de texture fine (argileuse), d'une pente forte (10-20%), et d'un sol peu profond (10 à 80 cm), d'un couvert végétal faible (Arboculture), Alors, elle présente un risque fort d'érosion.

Cela permet de doter l'ontologie d'un ensemble de règles liées à un domaine d'application donné à savoir l'érosion, l'inondation, et la déforestation. En effet, cela permet à l'ontologie de favoriser les mécanismes de raisonnement inférentiel sur le domaine.

2.4. Formalisme de l'ontologie proposée

Nous proposons le formalisme de l'ontologie suivant :

$O = (C, P, R, A, \leq^C)$ où :

- C est l'ensemble des concepts définis pour un domaine
- $P = (c, \mu_{\alpha,s,b,t}, \pi)$ où
 - $c \in C$,
 - $\mu_{\alpha,s,b,t}$ désigne le degré d'appartenance à un attribut α d'un concept c pour une bande spectrale b d'un capteur s à une date t
 - π est un poids attribué à la propriété.
- R est l'ensemble de relations spatiales (r, μ_r) où
 - r : type de la relation spatiale
 - μ_r désigne le degré d'appartenance à la relation spatiale floue
- $\leq^C : C \times C$ est un ordre partiel sur C , il définit la hiérarchie de concepts, $\leq^C(c_1, c_2)$ signifie que c_1 subsume c_2 (relation orientée)
- A est l'ensemble d'axiomes exprimés dans un langage logique flou

3. Présentation de l'ontologie résultante

Nous présentons dans cette section l'ontologie résultante suite à l'application du processus de développement proposé dans la section précédente.

3.1. Choix de ressources ontologiques candidates

Nous rappelons que nous avons choisi de réutiliser l'ontologie de Durand et al. (2007) et de la considérer comme ontologie de noyau vue sa capacité à modéliser les spécificités spatiales et spectrales des objets urbains, détectés dans une scène satellitaire.

Concernant le choix des ressources qui servent à enrichir l'ontologie de noyau au niveau conceptuel, nous avons retenu les ontologies suivantes (cités dans le chapitre précédent) :

OntoGeo-V2-47 : Le choix de cette ontologie est justifié par sa capacité de modéliser certains aspects liés à l'imagerie satellitaire tels que les objets naturels (reliefs, terrains, mers, segments d'eau, etc.), les zones inflammables, les zones d'inondation, les structures géologiques, les zones dangereuses, les substances dangereuses, les réacteurs nucléaires, les objets vulnérables, et les voies d'eau.

En outre, OntoGeo modélise un ensemble d'événements que peuvent subir ces objets, et décrit les données qui leur sont rattachées, à savoir : des métadonnées, des données spatio-temporelles, des données de localisation et des données thématiques.

DAFOE : Nous avons opté pour l'utilisation de cette ontologie pour sa couverture de plusieurs types d'objets géographiques, et pour sa richesse en termes de relations spatiales.

Theme : Cette ontologie présente l'avantage de modéliser les objets liés à l'érosion des côtes tels que la pluie, les fleuves, le vent, le niveau de la mer, les inondations, l'infrastructure, les caractéristiques océaniques et terrestres, etc.

FusionTopoCarto2: L'importance de cette ontologie réside dans le fait qu'elle est riche en concepts, et qu'elle modélise des entités hydrographiques, topographiques et agricoles.

FTT-01 : Elle modélise beaucoup de concepts géographiques qui couvrent les régions, les secteurs administratifs, les colis de terre, les caractéristiques hydrographiques et physio-géographiques.

3.2. Enrichissement conceptuel de l'ontologie de noyau

Nous essayerons dans cette partie, de présenter pour chacune des ontologies sources utilisées, les différentes correspondances trouvées avec l'ontologie de noyau. Dans le but de vérifier la similarité du voisinage direct pour chaque couple de concepts appariés, nous procédons à ajouter, pour chacun de ces couples, la mesure structurelle qui correspond. Les mappings établis sont de deux types : sémantiques et terminologiques.

A ce titre, nous mentionnons que les résultats que nous détaillerons ci-après, sont obtenus suite à l'application d'un seuil de similarité structurelle de 50% et d'un seuil de similarité sémantique de 30%. La sélection des mappings terminologiques dépend uniquement de l'apparition ou non des mots constituant le nom d'un concept cible dans le nom de son subordonné relatif à l'ontologie source.

3.2.1. Résultats d'alignement et d'enrichissement pour l'ontologie source *FTT.owl*

Nous présentons ci-après un résumé sur l'ensemble des mappings d'équivalence sémantiques et terminologiques trouvés entre les concepts de l'ontologie cible et leurs correspondants dans l'ontologie source *FTT.owl*. En effet, les calculs ont abouti à un nombre total de 30 équivalences, dont une est sémantique. Ces résultats trouvés prouvent la proximité des deux ontologies appariées, notamment dans les domaines de constructions (*Temples, ponds, mansion, roads, tramways, houses, hospitals, mosques, airports, runways, towers, cemeteries, etc.*), d'hydrographie (*rivers, oceans, seas, etc.*), et de végétation (*vegetation, trees, forests, deserts, etc.*). Nous constatons, en outre, qu'une quinzaine de ces correspondances, soit 50% du nombre total, est dotée d'une similarité structurelle supérieure à 80%, et ce contre six mappings, dont la similarité ne dépasse pas 20%. Ceci s'explique par la couverture de l'ontologie *FTT* à beaucoup de concepts communs à l'ontologie cible. Ces concepts modélisent les régions, les secteurs administratifs, les caractéristiques hydrographiques, etc. Cet important nombre d'équivalences renforce le choix de cette ontologie comme une source d'enrichissement, et aide à établir un maximum de relation de subsumption, entre les fils de chacun des concepts de l'ontologie source et son équivalent sémantique ou terminologique dans l'ontologie cible. Le tableau suivant résume l'ensemble de ces relations qui sont déterminées soit sémantiquement au moyen de la ressource lexicale *WordNet*, soit terminologiquement à travers les techniques d'inclusion de labels et de la distance d'édition, ou encore à travers la navigation dans les

descendants des concepts sources reliés par une relation d'équivalence avec les concepts cibles.

Itération : 1		
Ontologie source : FTT.owl		
Type de relation : Equivalence		
Technique d'alignement : sémantique		
Numéro	Concept source, Concept cible	Similarité structurelle
1	Woods, Forest	0,36
Technique d'alignement : Terminologique		
Numéro	Concept source, Concept cible	Similarité structurelle
1	Temples, Temple	0,84
2	Tramways, Tramway	0,86
3	Vineyards, Vineyards	0,92
4	Ponds, Pond	0,78
5	Trees, Trees	0,86
6	Mansions, Mansion	0,86
7	Tundras, Tundra	0,42
8	Wetlands, Wetlands	0,10
9	Roads, Road	0,51
10	Cemeteries, Cemeteries	0,10
11	Seas, Sea	0,36
12	Towers, Tower	0,20
13	Hospitals, Hospital	0,88
14	Houses, House	0,82
15	Deserts, Desert	0,12
16	Beaches, Beaches	0,30
17	Bridges, Bridge	0,14
18	Estuaries, Estuaries	0,92
19	Oceans, Ocean	0,27
20	Mosques, Mosque	0,84
21	Rivers, River	0,16
22	Badlands, Badlands	0,91
23	Glaciers, Glaciers	0,91
24	Dunes, Dunes	0,28
25	Pastures, Pastures	0,91
26	Vegetation, Vegetation	0,36
27	Airports, Airports	0,91
28	Forests, Forest	0,47
29	Runways, Runway	0,84

Tableau 4.3 : Liste de mappings d'équivalence pour l'ontologie source FTT.owl

Itération : 1		
Ontologie source : FTT.owl		
Type de relation : Subsumption		
Technique d'alignement : sémantique		
Numéro	Concept source isA Concept cible	Similarité structurelle
1	Laguna, Lake	0,47
2	Woods, Vegetation	0,36
Technique d'alignement : Terminologique		
Numéro	Concept source isA Concept cible	Similarité structurelle
1	Laguna, Lake	0,47
2	woods, Vegetation	0,36

3	trees , Forest	0,38
4	pine_groves , Forest	0,45
5	palm_groves , Forest	0,36
6	olive_groves , Fores	0,58
7	groves , Forest	0,31
8	coconut_groves , Forest	0,50
9	Arctic_land , Tundra	0,50
10	peat_cutting_areas , Wetlands	0,35
11	peatlands , Wetlands	0,42
12	swamps , Wetlands	0,47
13	intermittent_wetlands , Wetlands	0,57
14	backwaters , Wetlands	0,51
15	tidal_flats , Wetlands	0,61
16	riparian_areas , Wetlands	0,51
17	sloughs , Wetlands	0,30
18	marshes , Wetlands	0,33
19	mangrove_swamps , Wetlands	0,47
20	salt_marshes , Wetlands	0,52
21	bayous , Wetlands	0,45
22	quagmires , Wetlands	0,34
23	forested_wetlands , Wetlands	0,59
24	tombs , Cemeteries	0,43
25	sepulchers , Cemeteries	0,59
26	mausoleums , Cemeteries	0,42
27	graveyards , Cemeteries	0,42
28	grave_sites , Cemeteries	0,54
29	burying_grounds , Cemeteries	0,50
30	pagodas , Tower	0,44
31	fire_lookouts , Tower	0,35
32	carillons , Tower	0,43
33	stony_deserts , Desert	0,49
34	sandy_areas , Desert	0,50
35	arid_regions , Desert	0,33
36	strands , Beaches	0,52
37	viaducts , Bridge	0,52
38	trestles , Bridge	0,52
39	overpasses , Bridge	0,34
40	covered_bridges , Bridge	0,45
41	ocean_regions , Ocean	0,35
42	ocean_currents , Ocean	0,34
43	rios , River	0,48
44	rapids , River	0,45
45	bends_OPENBRACKETriverCLOSEBRACKET , River	0,36
46	interdune_troughs , Dunes	0,50
47	rain_forests , Forest	0,57
48	petrified_forests , Forest	0,59
49	national_forests , Forest	0,53
50	forest_reserves , Forest	0,59

Tableau 4.4 : Liste de mappings de subsomption pour l'ontologie source FTT.owl

D'après ces résultats (cf. Tableau 4.4), nous remarquons qu'un nombre limité de mappings (soit 9 parmi 50) intéresse les constructions (tours et cemeteries). Les concepts subsomants apparaissant dans ces mappings ne semblent pas prioritaires pour l'étude des risques naturels. Cependant, nous apercevons que le reste de mappings se répartit en deux grandes familles : l'hydrographie et la végétation. Le développement de chacune de ces deux familles semble utile pour l'étude de ces risques. En effet, les phénomènes d'inondation et d'érosion sont étroitement liés à la

nature des objets hydrographiques et de végétation, la nature du sol, et aux relations spatiales topologiques et de distance reliant ces différents objets.

La majorité de ces mappings est dotée d'une similarité structurelle proche de 50%. Ceci prouve la cohérence et l'importance des concepts qu'elle contient dans le domaine des images satellitaires. L'utilisation de cette même ontologie dans une deuxième itération permet d'extraire d'autres concepts et d'augmenter le niveau de détail de l'ontologie cible dans un domaine particulier. Le tableau 4 résume les résultats relatifs aux quatre premières itérations.

	Itération 1	Itération 2	Itération 3	Itération 4
Similarité globale	49%	53%	54%	55%
Equivalences sémantiques	1	1	1	1
Equivalences terminologiques	29	77	86	91
Subsommptions sémantiques	8	0	0	0
Subsommptions terminologiques	42	9	5	1
Nombre de concepts ajoutés	50	9	5	1

Tableau 4.5 : Résultats d'alignement pour quatre itérations sur l'ontologie FTT.owl

Au début de l'itération 1, le calcul de la similarité globale entre les deux ontologies nous donne un taux de 49%. Ceci peut s'expliquer, d'une part, par le nombre important de concepts communs aux deux ontologies (29 correspondances terminologiques et 1 correspondance sémantique), et d'autre part, par la ressemblance du voisinage pour une grande partie de concepts des deux ontologies alignées. Cette valeur élevée de la similarité globale renforce le choix de cette ontologie comme ontologie source, et confirme son intérêt à enrichir l'ontologie noyau.

Comme nous le remarquons (cf. tableau 4.5), l'itération 1 a permis d'ajouter 50 liaisons de subsomption à l'ontologie cible, dont 20 (soit 40%) sont dotées d'une mesure structurelle supérieure à 50%. Toutes ces liaisons sont considérées par l'expert comme significatives et cohérentes. La recherche de mappings dans chaque itération se limite à la découverte des fils directs pour leurs correspondants dans l'ontologie cible. Ceci nous permet de conduire le niveau de détail que nous voulons assurer pour l'ontologie cible, dans un domaine particulier. L'itération de cette étape nous permet de découvrir des concepts qui n'ont pas pu être détectés dans des itérations antérieures, et notamment d'augmenter la profondeur de l'ontologie puisqu'elles amènent à l'ajout de nouveaux fils à des concepts qui ont été ajoutés dans les itérations précédentes.

La deuxième itération nous a apporté 9 nouveaux concepts. Parmi ces concepts, nous distinguons *fossilized_forests* et *marine_regions* fils respectifs des concepts *petrified_forests* et *ocean_regions* qui ont été déjà découverts dans l'itération 1. Les autres concepts sont ajoutés à des concepts qui existent dans l'ontologie dans son état initial. Leur apparition dans cette itération revient à l'augmentation de leur

mesure structurelle. Ces concepts sont *sandy_deserts* et *rock_deserts* fils de *Desert*, *waterfalls* fils de *River*, *memorial_gardens* et *graves* fils de *Cemeteries*, *cienagas* et *mires* fils de *Wetlands*. remarquons que les concepts *Desert*, *River*, *Cemeteries*, et *Wetlands* existent dans l'ontologie cible, dans son état initial. Ces concepts ont été enrichis, dans l'itération 1, par de nouveaux fils, ce qui signifie l'évolution de leur voisinage, et par conséquent de leur proximité aux concepts de l'ontologie source. Ceci mènera à la découverte de nouvelles correspondances dans des itérations ultérieures.

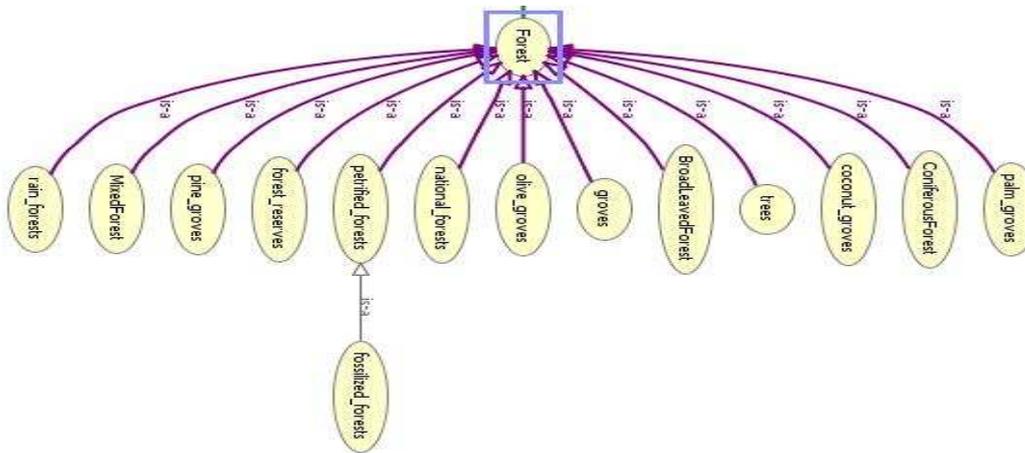


Figure 3.10 : Extrait de l'ontologie résultante après l'ajout des descendants du concept *Forest*

L'itération 3 a mené à cinq nouveaux concepts qui sont *bajadas* fils de *Desert*, *falls*, *cataracts*, *cascades* fils de *waterfalls*, et *fens* fils de *Wetlands*. Le concept *waterfalls* est à son tour ajouté dans l'itération 2. Les résultats de cette itération montrent que le nombre de mappings découverts diminue d'une itération à une autre. Cette constatation est soutenue par l'itération 4 qui n'a permis que la découverte d'une seule nouvelle correspondance qui est : *slues* isA *Wetlands*.

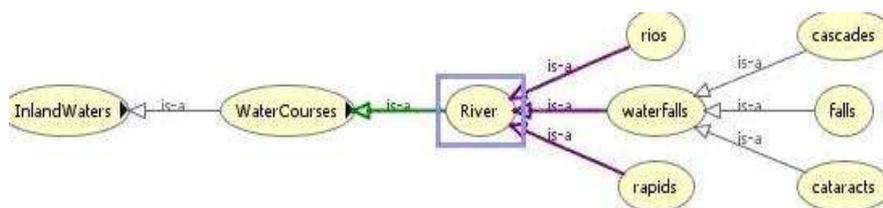


Figure 3.11 : Extrait de l'ontologie résultante après l'ajout des descendants du concept *waterfalls*

3.2.2. Résultats d'alignement et d'enrichissement pour l'ontologie source hydrology.owl

Nous détaillons dans cette partie, les résultats d'alignement obtenus suite à l'utilisation de l'ontologie hydrology.owl. Le but étant d'enrichir l'ontologie cible par

les concepts liés à l'hydrographie à cause de leur importance dans les prévisions des risques d'inondation et d'érosion.

Les résultats obtenus après l'application des seuils donnés auparavant sont les suivants :

Itération : 1		
Ontologie source : Hydrology.owl		
Type de relation : Equivalence		
Technique d'alignement : Sémantique		
Numéro	Concept source, Concept cible	Similarité structurelle
1	Pool, Pond	0,25
2	Waterfall, falls	0,78
Technique d'alignement : Terminologique		
Numéro	Concept source, Concept cible	Similarité structurelle
1	Lake, Lake	0,32
2	Cataract, cataracts	0,88
3	Canal, Canal	0,5
4	Pond, Pond	0,1
5	Pasture, Pastures	0,98
6	River, River	0,57
7	Rapid, rapids	0,82
8	Fen, fens	0,72
9	Sea, Sea	1,00
10	Cascade, cascades	0,81
11	Swamp, swamps	0,82
12	Backwater, backwaters	0,89
13	Wetland, Wetlands	0,60
14	Road, Road	0,50

Tableau 4.6 : Liste de mappings d'équivalence pour l'ontologie source Hydrology.owl

Itération : 1		
Ontologie source : Hydrology.owl		
Type de relation : Subsumption		
Technique d'alignement : Sémantique		
Numéro	Concept source <i>isA</i> Concept cible	Similarité structurelle
1	Tarn, Lake	0,55
2	Castle, Mansion	0,53
3	Cascade, falls	0,56
4	Pot, Pond	0,77
5	PlungePool, Pond	0,61
6	Cataract, falls	0,44
7	DecoyPond, Pond	0,45
8	OxbowLake, Pond	0,45
9	ModelBoatLake, Pond	0,55
Technique d'alignement : Terminologique		
Numéro	Concept source <i>isA</i> Concept cible	Similarité structurelle
1	Llyn, Lake	0,58
2	Tarn, Lake	0,55
3	Loch, Lake	0,55
4	Broad, Lake	0,46

6	NavigationalCut, Canal	0,51
7	Marsh, Wetlands	0,42
8	CanalStretch, Canal	0,90
9	IrrigationCanal, Canal	0,42
10	CanalFeeder, Canal	0,90
11	BalancingPond, Pond	0,44
12	DecoyPond, Pond	0,45
13	RiverBed, River	0,37
14	RiverMouth, River	0,44
15	BraidedRiver, River	0,40
16	TidalRiver, River	0,37
17	BraidedRiverStretch, River	0,30
18	RiverStretch, River	0,42
19	TidalRiverStretch, River	0,35
	RiverBank, River	0,47

Tableau 4.7 : Liste de mappings de subsomption pour l'ontologie source *Hydrology.owl*

Les résultats, décrits dans le tableau 4.5, montrent que plus que 80% des mappings d'équivalence sémantiques et terminologiques sont dotés d'une mesure structurelle qui dépasse 50%. Ces valeurs révèlent, d'un côté, la proximité de l'ontologie *Hydrology.owl* au domaine de l'imagerie satellitaire, et d'un autre côté, son importance pour enrichir les concepts de l'ontologie cible.

Les relations de subsomption obtenues (cf. tableau 4.7) renforcent cette constatation. En effet, nous constatons que 66% des mappings sémantiques disposent d'une mesure structurelle supérieure à 50%, et que 30% des mappings terminologiques sont dotés d'une mesure structurelle qui dépasse 50%. Cependant, il est important de remarquer que les mappings sémantiques sont jugés par l'expert comme intéressants pour l'enrichissement de l'ontologie cible malgré que certains d'entre eux ont une similarité structurelle faible (inférieure à 0.4). Ceci peut s'expliquer par la pauvreté de l'ontologie cible en termes de concepts liés à l'hydrographie, ce qui signifie qu'un concept cible n'a pas beaucoup de voisins en commun avec ses correspondants dans l'ontologie source. Le maintien de ces mappings est justifié par la proximité élevée des deux ontologies appariées, ce qui permet d'éviter les éventuels problèmes d'homonymie. Par conséquent, toute relation sémantique prouvée par WordNet est jugée consistante.

Pour les mappings terminologiques de subsomption (cf. tableau 4.7), nous remarquons que notre prototype a abouti à un nombre important de correspondances, dont 13 (soit 68% du nombre total des mappings) ont une similarité structurelle inférieure à 50%. Toutefois, en essayant d'analyser leurs relations par rapport au domaine de l'hydrographie, nous remarquons que la majorité d'entre eux sont significatifs. Le facteur essentiel justifiant la faiblesse structurelle d'une grande

partie de ces mappings est la limitation de la hiérarchie des concepts liés à l'hydrographie dans l'ontologie cible.

D'un autre côté, nous nous apercevons que les concepts découverts sont répartis sur quatre familles, décrites par les concepts nommés *Lake*, *Canal*, *Pond* et *River*. Ceci montre que l'extension de l'ontologie cible, dans le domaine de l'hydrographie, est orientée par les concepts hydrographiques qu'elle contient dans son état initial.

Le tableau 4.8 illustre les résultats obtenus pour chacune des quatre premières itérations exécutées sur l'ontologie source *Hydrology.owl*.

	Itération 1	Itération 2	Itération 3	Itération 4
Similarité globale	58%	60%	60,50%	60,80%
Equivalences sémantiques	2	2	2	2
Equivalences terminologiques	14	38	43	46
Subsommptions sémantiques	10	1	2	0
Subsommptions terminologiques	19	12	6	1
Nombre de concepts ajoutés	26	7	4	1

Tableau 4.8 : Résultats d'alignement pour quatre itérations sur l'ontologie *Hydrology.owl*

La première itération (cf. tableau 4.8) a mené à 26 nouveaux concepts. A cet égard, nous remarquons que le nombre de concepts trouvés dans les mappings de subsomption ne correspond pas au nombre total des concepts ajoutés. Ceci est dû à la présence d'un certain nombre de redondances entre les résultats obtenus pour chacune des deux techniques utilisées. Les mappings redondants générés dans la première itération sont : *Tarn isA Lake*, *Cascade isA falls* et *DecoyPond isA Pond*. En effet, les concepts *Tarn* et *Cascade* apparaissent comme les fils respectifs de *Lake* et de *falls* dans deux situations, une première, au moyen d'une relation d'hyponymie qui relie l'identificateur de chacun d'eux, à celui de son correspondant dans l'ontologie cible, et une deuxième, comme étant les fils directs de leurs équivalents dans l'ontologie cible. Quant au concept *decoyPond*, il apparaît une première fois dans une relation terminologique, vérifiant ainsi la propriété d'inclusion de labels avec son correspondant dans l'ontologie cible, et une deuxième fois, comme étant le fils direct du concept *Pool*, généré comme équivalent au concept *Pond* de l'ontologie cible.

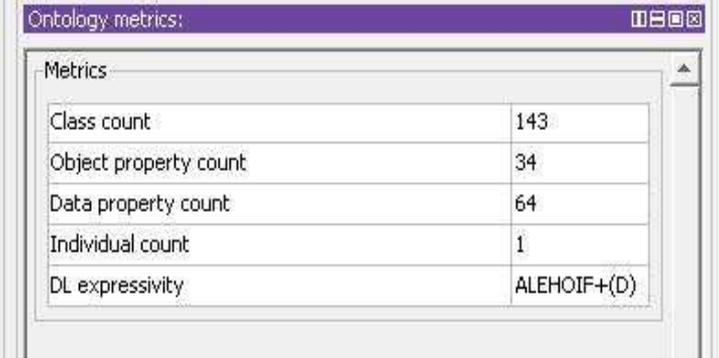
D'un autre côté, nous remarquons que cette itération a conduit à 11 mappings de subsomption (soit 44%) qui sont dotées d'une mesure structurelle supérieure à 50%. Ce taux n'a cessé d'augmenter dans les itérations suivantes. En effet, il a évolué à 48 % dans la deuxième itération et à 56% dans la troisième itération.

Enfin, nous ajoutons que l'utilisation de l'ontologie Hydrology.owl, nous a permis d'enrichir notre ontologie de départ par 38 nouveaux concepts, repartis sur deux niveaux hiérarchiques.

3.2.3. Résultats finaux

Dans ce qui suit nous présentons des aperçus sur les résultats trouvés suite à la phase de placement de concepts. Nous commençons par présenter un certain nombre de métriques se rapportant à l'ontologie cible. Ces métriques résument l'état de l'ontologie (i) dans son état initial, (ii) et après son enrichissement par de nouveaux concepts.

Etat initial : L'ontologie de départ, considérée comme le noyau de l'ontologie finale, renferme 143 concepts, 34 propriétés d'objets, 64 propriétés de données et un seul individu.



Metrics	
Class count	143
Object property count	34
Data property count	64
Individual count	1
DL expressivity	ALEHOIF+(D)

Figure 3.12 : Quelques métriques sur l'ontologie de départ

La figure suivante montre l'extension de l'ontologie cible par des nouvelles classes liées au domaine hydrographique. Ces concepts sont ajoutés sur plusieurs niveaux hiérarchiques, augmentant ainsi la hauteur de l'ontologie résultante.



Figure 3.13 : Extrait de l'ontologie résultante après Ajout du concept waterfalls et ses fils

Le nouvel état de l'ontologie après l'exécution de quatre itérations sur chacune des ontologies sources FTT.owl et Hydrology.owl, est résumé dans la figure suivante :

Metrics	
Class count	242
Object property count	34
Data property count	64
Individual count	1
DL expressivity	ALEHOIF+(D)

Figure 3.14 : Quelques métriques sur l'ontologie après son enrichissement

3.3. Enrichissement intensionnel de l'ontologie de noyau

L'enrichissement intensionnel de l'ontologie de noyau se base essentiellement sur l'information de bas niveau extraite de l'image de télédétection.

3.3.1. Elaboration d'une banque d'imagettes

Nous avons conçu une banque d'imagettes multi-dates et multi-capteurs représentant des classes d'ontologie (forêt, lac, champ de blé, zone urbaine, ...), en

utilisant des outils de traitement d'images satellitaires à savoir ENVI®¹, ERDAS IMAGINE®², et VISAT®³. Nous avons utilisé les bases d'images LandSat TM (7 bandes) fournies gratuitement par le site USGS⁴ représentant la région du Kef de la Tunisie, et des images SPOT représentant la même zone fournies par le Ministère d'agriculture Tunisien. La figure 3.15 illustre un extrait de la banque montrant des imagettes de forêts et de lac issues de deux capteurs LandSat et SPOT à deux dates.

3.3.2. Illustration du processus d'enrichissement intensionnel

La figure 3.16 illustre le processus d'élaboration des fonctions d'appartenance gaussiennes représentant les propriétés des concepts de l'ontologie en fonction de temps t et du capteur s .

- 1) Extraction de la région d'intérêt de l'imagette avec l'algorithme de segmentation ISODATA (Ball et hall, 1965);
- 2) Extraction des primitives radiométriques (spectrales et indices : INDVI et MNDWI), de textures, de formes de chaque région ;
- 3) Représentation gaussienne des primitives de chaque région, avec le logiciel Matlab ;
- 4) Alimentation de l'ontologie par les valeurs *a priori* floues extraites de l'image;

¹ <http://www.exelisvis.com/>

² <http://www.geosystems.fr/>

³ <http://www.brockmann-consult.de/cms/web/beam/>

⁴ <http://www.usgs.com>

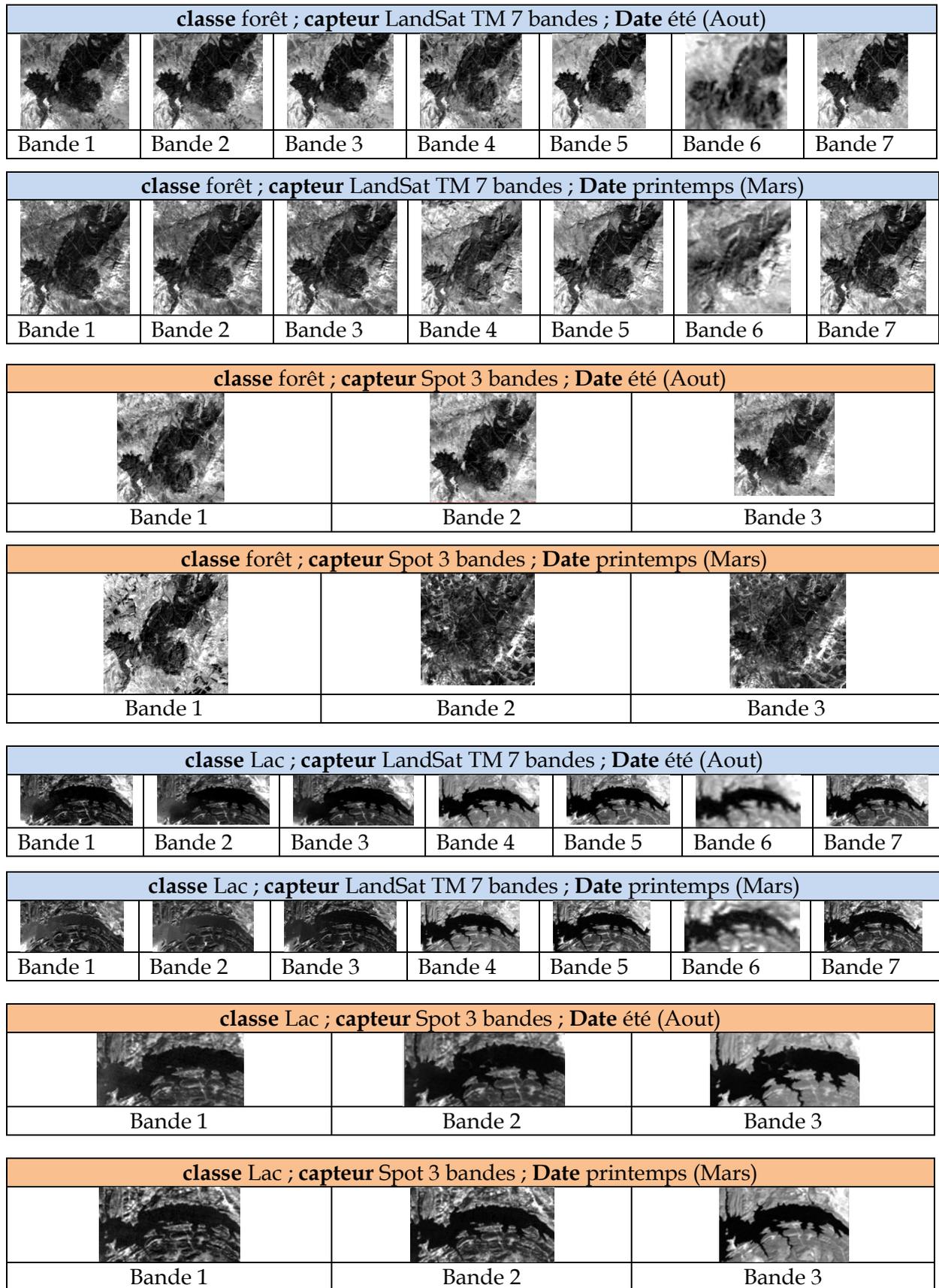


Figure 3.15 : Extrait de la banque d'images

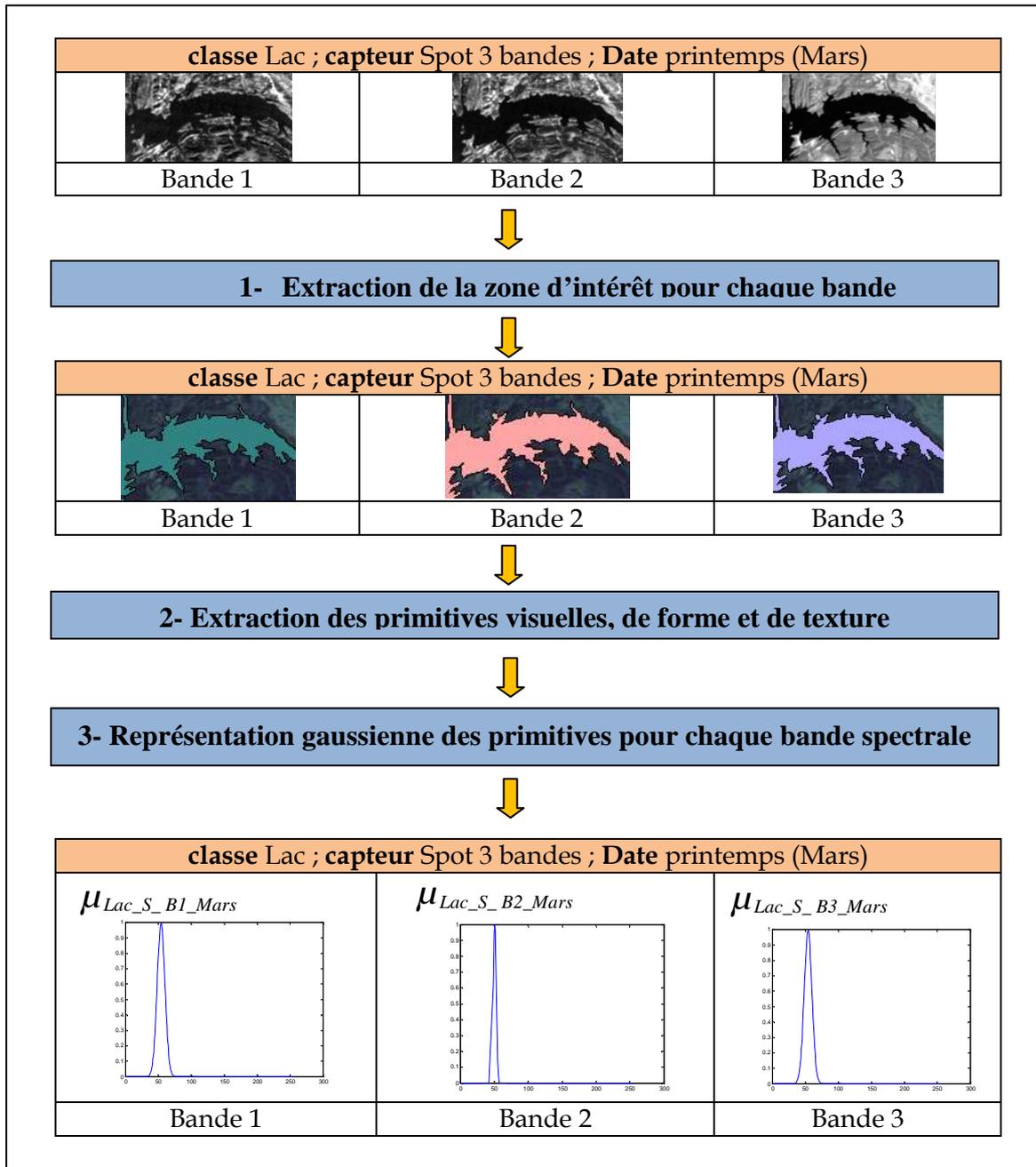


Figure 3.16 : *Processus d'élaboration des fonctions d'appartenances*

3.3.3. Structure intensionnelle d'un concept

Le tableau 4.9 illustre la structure du concept « lac » défini à partir des fonctions d'appartenances gaussiennes représentant des primitives visuelles (radiométrie, texture et forme) selon les capteurs SPOT et LandSat, à deux dates (printemps et été).

Classe « Lac »				
Primitives	SPOT(S)		LANDSAT(L)	
	Eté(E)	Printemps(P)	Eté(E)	Printemps(P)
Radiométrie	$\mu_{R_S_B1_E}$ $\mu_{R_S_B2_E}$ $\mu_{R_S_B3_E}$...	$\mu_{R_S_B1_P}$ $\mu_{R_S_B2_P}$ $\mu_{R_S_B3_P}$...	$\mu_{R_L_B1_E}$ $\mu_{R_L_B2_E}$ $\mu_{R_L_B3_E}$... $\mu_{R_L_B7_E}$	$\mu_{R_S_B1_P}$ $\mu_{R_S_B2_P}$ $\mu_{R_S_B3_P}$... $\mu_{R_S_B7_P}$
NDVI	$\mu_{NDVI_S_E}$	$\mu_{NDVI_S_P}$	$\mu_{NDVI_L_E}$	$\mu_{NDVI_L_P}$
MNDWI	$\mu_{MNDWI_S_E}$	$\mu_{MNDWI_S_P}$	$\mu_{MNDWI_L_E}$	$\mu_{MNDWI_L_P}$
Texture	$\mu_{Tex_S_B1_E}$ $\mu_{Tex_S_B2_E}$ $\mu_{Tex_S_B3_E}$...	$\mu_{Tex_S_B1_P}$ $\mu_{Tex_S_B2_P}$ $\mu_{Tex_S_B3_P}$...	$\mu_{Tex_L_B1_E}$ $\mu_{Tex_L_B2_E}$ $\mu_{Tex_L_B3_E}$... $\mu_{Tex_L_B7_E}$	$\mu_{Tex_L_B1_P}$ $\mu_{Tex_L_B2_P}$ $\mu_{Tex_L_B3_P}$... $\mu_{Tex_L_B7_P}$
Forme_surface	$\mu_{Sur_S_E}$	$\mu_{Sur_S_P}$	$\mu_{Sur_L_E}$	$\mu_{Sur_L_P}$
Forme_périmètre	$\mu_{Per_S_E}$	$\mu_{Per_S_P}$	$\mu_{Per_L_E}$	$\mu_{Per_L_P}$

Tableau 4.9 : Structure intensionnelle du concept « lac »

3.3.4. Enrichissement axiomatique

Cet enrichissement permet de doter l'ontologie des mécanismes de raisonnement inférentiel. Nous présentons ci-après un extrait de la base de règles d'inférences ajoutées à l'ontologie :

R1 : Si un champ de blé est situé à coté d'une montagne de pente supérieure à 10°, alors le risque d'érosion est Elevé.

R2 : Si un champ de blé est situé à coté d'une montagne de pente inférieure à 1°, alors le risque d'érosion est Faible.

R3 : Si un champ de blé est situé à coté d'une rivière et le type de sol est argileux, alors le risque d'érosion est élevé.

Conclusion

Nous avons proposé, dans ce chapitre, de suivre un processus de développement d'ontologie basé sur la réutilisation des ressources ontologiques existantes. Ce processus consiste à choisir une ontologie de noyau, de l'adapter aux spécificités des images de télédétection, et de l'enrichir par des nouveaux concepts issus des ressources existantes. En termes de résultats, nous avons obtenu une ontologie spatio-temporelle, floue, lourde et évolutive, enrichi par des informations issues des images de télédétection et des axiomes modélisant les connaissances de domaine.

Ce chapitre présente la contribution concernant la proposition d'une méthodologie d'interprétation sémantique d'image de télédétection, guidée par l'ontologie développée, permettant une annotation sémantique floue de l'image de télédétection, et contribuant à déterminer la susceptibilité d'une scène à un risque naturel donné.

1. Proposition d'une méthodologie d'interprétation sémantique d'images de télédétection

Nous proposons une méthodologie permettant de générer des annotations modélisant le contenu sémantique d'une image satellitale en se guidant par des connaissances de domaine représentées par l'ontologie proposée. Aussi, à partir d'un raisonnement inférentiel sur ces connaissances, nous pouvons déduire la susceptibilité d'une scène satellitale aux risques naturels. La méthodologie proposée comporte quatre phases : (1) Modélisation de la scène, (2) Fusion des annotations, (3) Recherche de cas, et (4) Raisonnement et interprétation (Figure 4.1).

Les objectifs attendus de notre méthodologie sont :

- Extraire les informations et les connaissances de l'image satellitale.
- Modéliser le contenu des images en se basant sur les connaissances de domaine, et en tenant compte du capteur utilisé pour l'acquisition et de la dimension temporelle.
- Analyser le contenu de l'image et déduire si la scène qu'elle représente présente un risque naturel, en se basant sur un raisonnement inférentiel basé sur les connaissances du domaine et la base de cas.

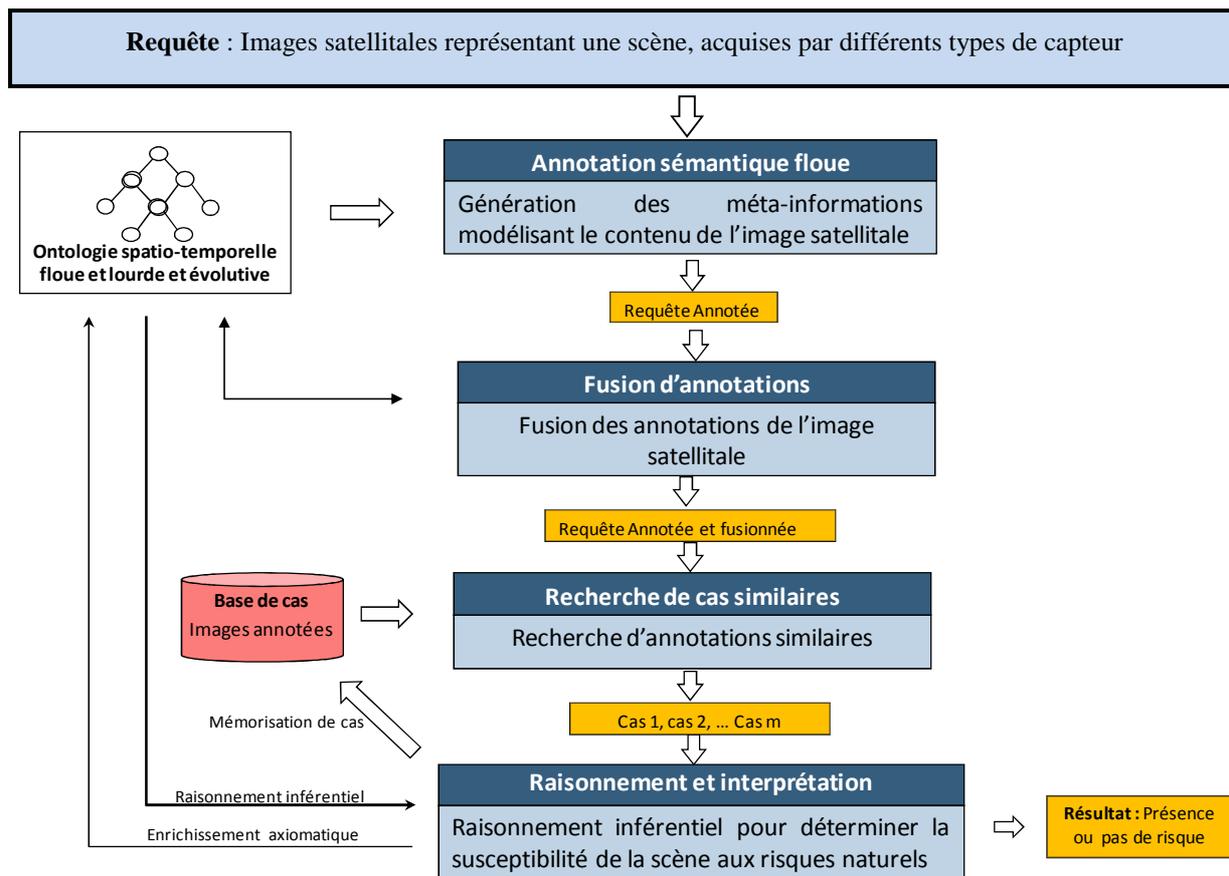


Figure 4.1: Méthodologie proposée pour l'interprétation sémantique d'images de télédétection

Nous présentons en détail dans les sections suivantes chaque phase de notre méthodologie.

2. Annotation floue de l'image de télédétection guidée par ontologie

L'objectif est de générer une annotation sémantique de l'image de télédétection décrivant son contenu sémantique en se basant sur l'ontologie proposée, tout en tenant compte des imperfections que peut avoir l'image. Nous proposons alors une annotation floue permettant de passer d'une information de bas niveau (décrite à partir des primitives de l'image) à une information de haut niveau (décrite par un modèle).

2.1. Processus de génération d'annotation

Ce processus que nous proposons permet d'établir la correspondance entre l'information issue de l'image d'un côté, et les concepts de l'ontologie, d'un autre côté, comme suit :

- 1) On fait l'extraction des régions de l'image de télédétection et leur primitives de bas niveau : radiométrie, texture, forme et indices.
- 2) On modélise les régions de l'image par un ensemble de sous ensembles flous modélisant ses primitives.
- 3) On cherche les concepts de l'ontologie les plus similaires à la région de l'image : Processus de *Matching flou* que l'on détaille dans la sous-section 2.2
- 4) Génération de l'annotation qui décrit le contenu sémantique de l'image.

Ce processus est illustré par la figure 4.2.

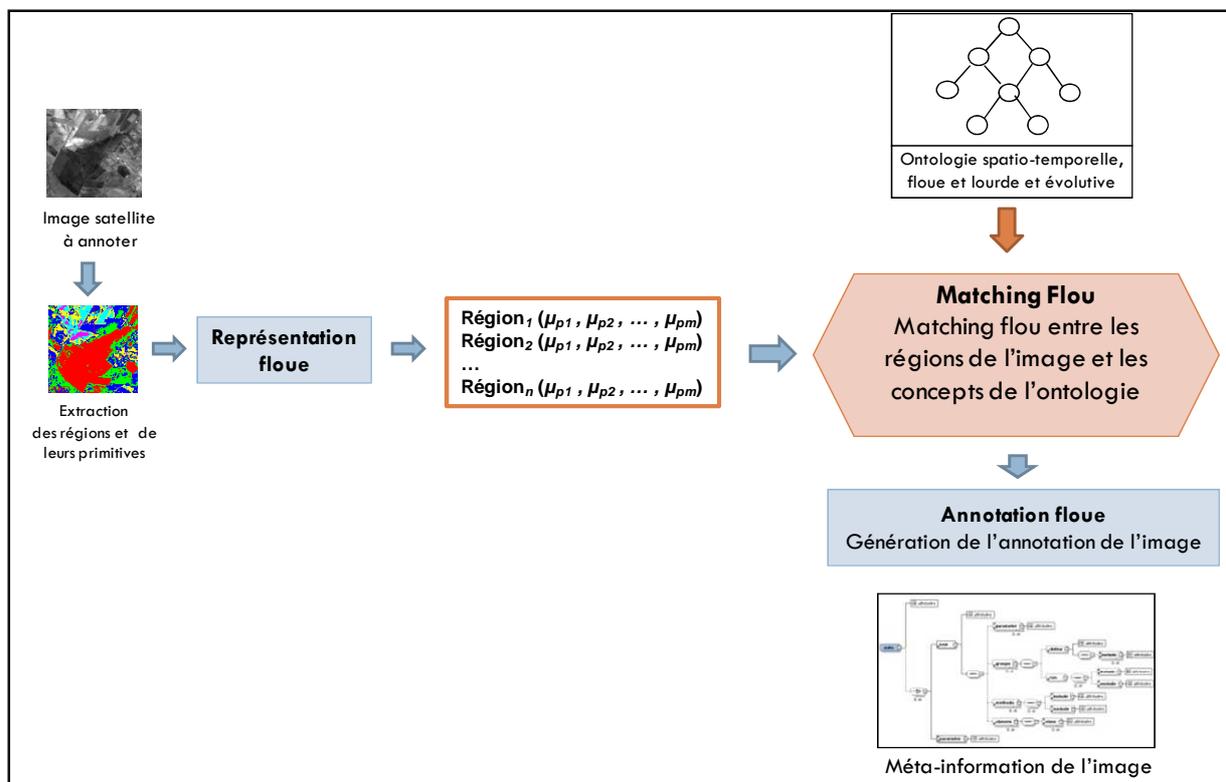


Figure 4.2 : *Processus de génération d'annotation d'image de télédétection*

2.2. *Processus de matching flou*

Ce processus permet de déterminer le ou les concepts de l'ontologie qui représente la ou les régions de l'image de télédétection. Cela se fait en mesurant la similarité entre les sous-ensembles flous de la région de l'image et celles des concepts de l'ontologie. Dans ce contexte, Johaniyak et al. (2005) a proposé une mesure de distance entre les sous ensembles flous basée sur les distances :

Soient A et B deux ensembles flous avec les fonctions d'appartenances respectives μ_A et μ_B , et $X=\{x_1, x_2, x_3, \dots, x_n\}$ des éléments de A et B ; la similarité entre A et B est définie :

$$S(A, B) = \frac{1}{1 + D(A, B)}$$

Avec $D(A,B)$: Distance euclidienne floue entre A et B :

$$D(A, B) = \sqrt{\sum_{i=1}^n (\mu_A(x_i) - \mu_B(x_i))^2}$$

La figure 4.3 illustre le processus du Matching flou proposé. On définit la similarité $S(\mathbf{R}_i, C_i)$ entre la région \mathbf{R} de l'image ayant un ensemble de primitives ($Pr_1, Pr_2, \dots Pr_n$) et le concept C de l'ontologie ayant un ensemble de propriétés ($Pc_1, Pc_2, \dots Pc_n$) comme suit :

$$S(R_i, C_i) = \frac{\sum_{i=1}^n Sp(Pr_i, Pc_i)}{n}$$

Avec

$$Sp(Pr_i, Pc_i) = \frac{1}{1 + D(Pr_i, Pc_i)}$$

Avec,

$$D(Pr_i, Pc_i) = \sqrt{\sum_{i=1}^n (\mu_{Pr_i}(x_i) - \mu_{Pc_i}(x_i))^2}$$

Nous calculons alors le score de similarité entre une région de l'image et chaque concept de l'ontologie, et nous retenons le concept ayant le score maximal. Ce processus est itéré pour toutes les régions de l'image de télédétection. L'ensemble des correspondances retenues constituent l'annotation de l'image que l'on détaillera dans la sous-section 2.3.

2.3. Méta-information proposée de l'image de télédétection

Le résultat du matching flou est un modèle ayant la structure suivante (Figure 4.4):

$A_{im}(O_{im}, R_{im}, C_{im}, Ct_{im})$ où :

- O_{im} est l'ensemble d'objets de l'image ; chaque objet a un score de similarité $S(R_i, C_i)$ avec le concept correspondant de l'ontologie .
- R_{im} est l'ensemble des relations spatiales entre les objets
- C_{im} est l'ensemble des caractéristiques de l'image
- Ct_{im} est l'ensemble de contraintes de l'image

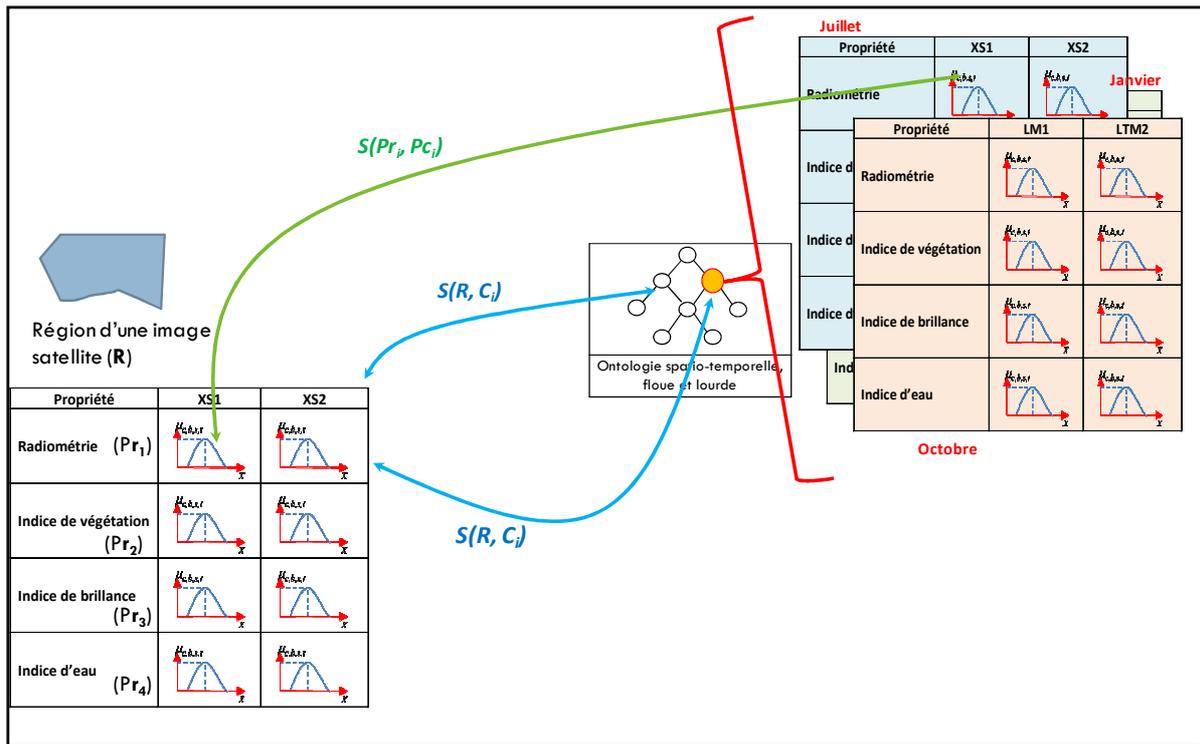


Figure 4.3 : Illustration du processus de matching flou

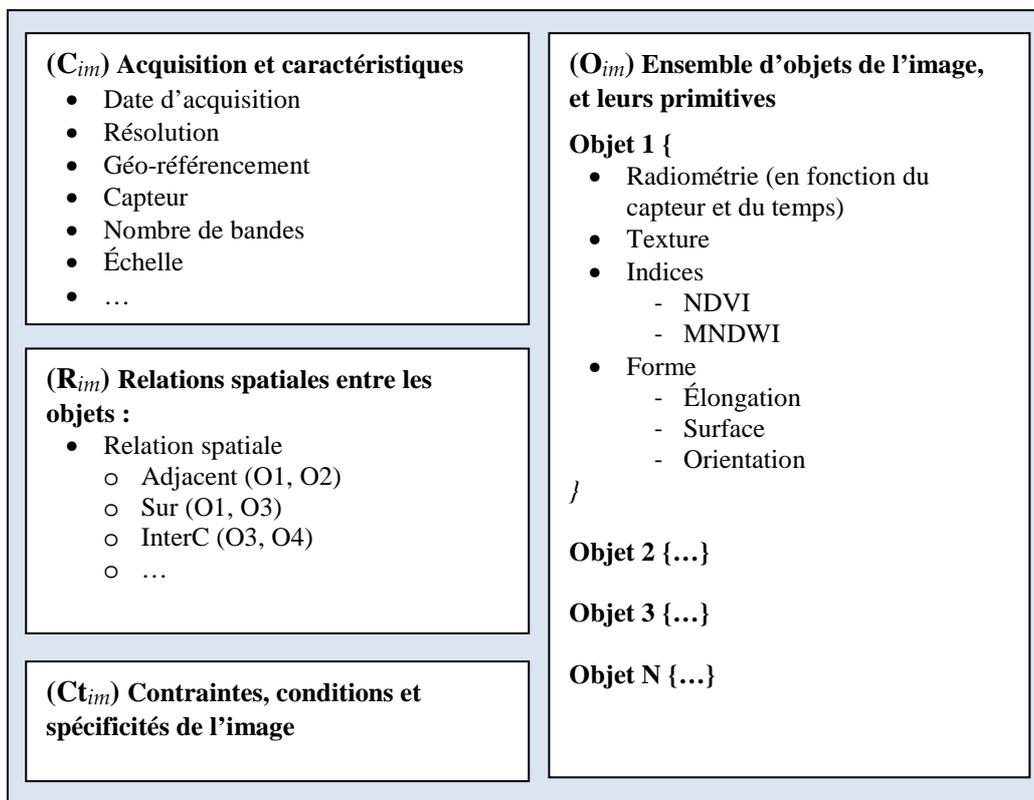


Figure 4.4 : Méta-information proposée de l'image de télédétection

La figure 4.5 présente une représentation XML générale d'une méta-information.

```

<Annotation>
  <Caractéristiques>
    <Capteur> ... </Capteur>
    <nb_Bandes>... </nb_Bandes>
    <Résolution> ... </Résolution>
    <Géo-Ref>
      <Proj> ... </Proj>
      <Datum> ... </Datum>
    </Géo-Ref>
  </Caractéristiques>
  <Objets>
    <Objet>
      <id_objet>... </id_objet>
      <concept_ontologie>
        <Nom_concept> ... </Nom_concept>
        <Degré_Appartenance> ... </Degré_Appartenance>
        <Primitives>
          <Radiométrie>
            <μR_S_B1_P>
              <x> ... </x>
              <y> ... </y>
            </μR_S_B1_P>
            <μR_S_B2_P> ... </μR_S_B2_P>
            ...
          </Radiométrie>
          <Texture> ... </Texture>
          <Forme> ... </Forme>
        </Primitives>
      </Objet>
    <Objet> ... </Objet>
    ...
  </Objets>
  <Relations_spatiales>
    <Relation_spatiale>
      <type>... </type>
      <Objet1> ... </Objet1>
      <Objet2> ... </Objet2>
      <Degré_Appartenance> ... </Degré_Appartenance>
    </Relation_spatiale>
    <Relation_spatiale> ... </Relation_spatiale>
    ...
  </Relations_spatiales>
  ...
</Annotation>

```

Figure 4.5 : Représentation XML d'une méta-information

3. Fusion floue des annotations d'images de télédétection

Afin d'avoir un maximum d'information, différents type de capteurs peuvent être utilisés pour l'acquisition d'images satellitales relatives à une même scène. Ces images, issues de différents capteurs, ont les spécificités suivantes :

- Un objet peut être plus clair dans une image radar que dans une image optique.
- Un objet peut avoir des significations différentes selon le type de capteur.

Dans le but de profiter au mieux de cette richesse d'information (vue la diversité des capteurs) et d'avoir une information fidèle à la réalité, et par conséquent, avoir des interprétations correctes sur la susceptibilité aux risques naturels d'une scène, nous proposons la démarche suivante (Figure 4.6) :

- 1) On modélise chaque image satellitale en appliquant le module précédent. On obtient ainsi plusieurs annotations représentant la même scène.
- 2) On fusionne ces annotations pour avoir une annotation plus représentative du contenu de l'image.

3.1. *Processus de fusion proposé*

Nous proposons un algorithme permettant de fusionner les annotations d'images multi-capteurs représentant la même scène. L'algorithme se base sur les scores de similarités associés avec les objets afin de comparer ces derniers, nous avons alors trois cas possibles :

Cas 1 : Objet qui existe dans une seule annotation : on les ajoute dans la nouvelle annotation (opération d'union).

Cas 2 : Objets qui existent dans plusieurs annotations avec même score de similarité : on ajoute l'objet dans la nouvelle annotation (opération d'intersection).

Cas 3 : Objets qui existent dans plusieurs annotations avec scores de similarité différents : on calcule le nouveau score de similarité en utilisant la fusion floue (les moyennes), puis on ajoute dans le nouveau modèle l'objet ayant le score maximal.

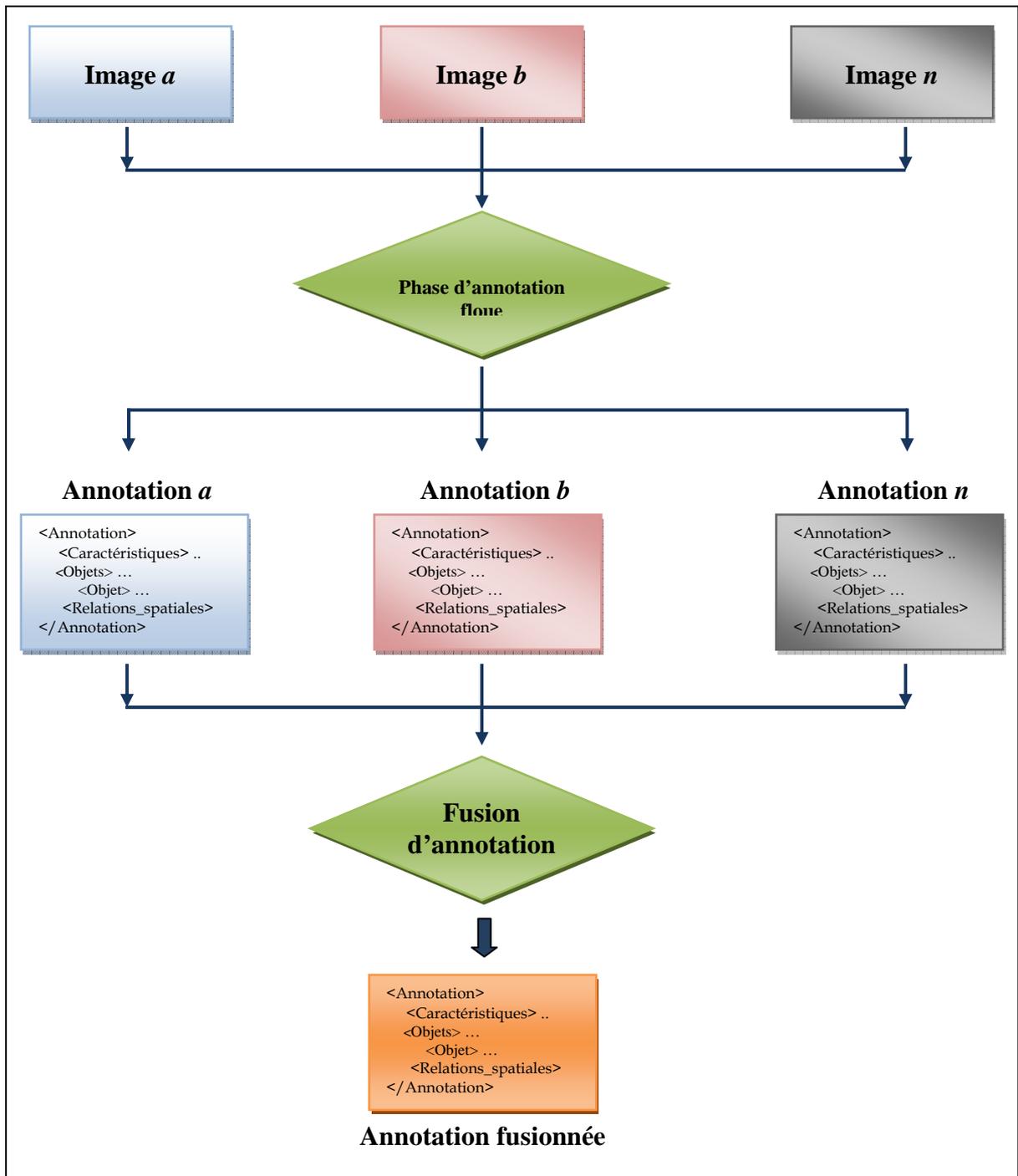


Figure 4.6 : *Processus de fusion d'annotations*

Algorithme de fusion d'annotations d'images:

Entrées

$A(O_A, R_A, C_A, C_{tA}), B(O_B, R_B, C_B, C_{tB})$: deux annotations représentant une même scène

Sorties

$M(O_M, R_M, C_M, C_{tM})$: Annotation résultante de la fusion de A_1 et A_2 .

Variables intermédiaires

objet : objet de l'ensemble O de l'annotation.

Fonctions

Ajouter($x, Y.O_Y$) : Ajoute l'objet x dans l'ensemble d'objets O_Y de l'annotation Y .

Moyenne_floue($objet, sc_1, sc_2$) : Calcule la moyenne floue entre sc_1 et sc_2 , et l'affecte à $objet$

Début

Itérer

Si (($objet$ existe dans $A.O_A$ et n'existe pas dans $B.O_2$) Ou ($objet$ existe dans $B.O_2$ et n'existe pas dans $A.O_1$)) Alors

Ajouter($objet, M$.)

Sinon // ($objet$ existe dans les deux annotations)

Si ($objet$ a le même degré d'appartenance dans les deux annotations A_1 et A_2) Alors

Ajouter($objet, M$)

Sinon // ($objet$ a des degrés d'appartenances différentes)

Moyenne_floue($objet, Score_objet_O_A, Score_objet_O_B$)

Ajouter($objet, M$)

Fin Si

Fin Si

Fin

Fin

4. Interprétation de l'image de télédétection à base d'ontologie et du RBC

4.1. Présentation du système d'interprétation proposé

La phase d'annotation sémantique permet de constituer une base d'images annotées par des méta-informations. Cette base représente le contenu des images de télédétection, leurs primitives visuelles, leurs propriétés spatio-temporelles, et leurs contraintes et spécificités.

Nous proposons d'exploiter la base d'annotation et de réutiliser l'ontologie de domaine pour déduire la susceptibilité d'une scène requête à un risque naturel donné, en se basant sur des mécanismes de raisonnement à base de cas.

Le raisonnement à base de cas consiste à résoudre un problème en se remémorant et en adaptant des cas passés déjà résolus. Les systèmes de RBC manipulent des connaissances de natures diverses : les cas, les connaissances du domaine, les connaissances de similarité et d'adaptation que nous les détaillons dans la section 4.4.

La figure 4.7 illustre notre système d'interprétation d'images de télédétection basé sur le raisonnement à base de cas et d'ontologie. Ce système est composé de :

La requête (le nouveau cas) : c'est le résultat du module de l'annotation sémantique. Il s'agit de la méta-information représentant le contenu sémantique d'une scène satellite requête. L'objectif est de déduire sa susceptibilité à un risque naturel donné.

La base de cas : il s'agit de l'ensemble des images de télédétection annotées et interprétées. L'alimentation de la base de cas correspond à la mémorisation des annotations des images et leurs interprétations.

L'ontologie : c'est l'ontologie de domaine proposée qui va guider le raisonnement et l'interprétation de l'image.

La recherche de cas similaires : c'est une phase permettant de rechercher les cas similaires à la requête, en se basant sur des mesures similarités entre les cas.

L'adaptation et l'ajustement : ces phases permettent de trouver des solutions pour le cas cible à partir des cas candidates. Elle se base sur des connaissances d'adaptation et celles issues de l'ontologie (voir section 1.3.2.3).

La mémorisation : consiste à mettre à jour les connaissances du système à l'issue du raisonnement, en mémorisant le nouveau cas dans la base de cas.

L'enrichissement : Cette phase permet d'enrichir l'ontologie de domaine par des nouvelles connaissances inférées.

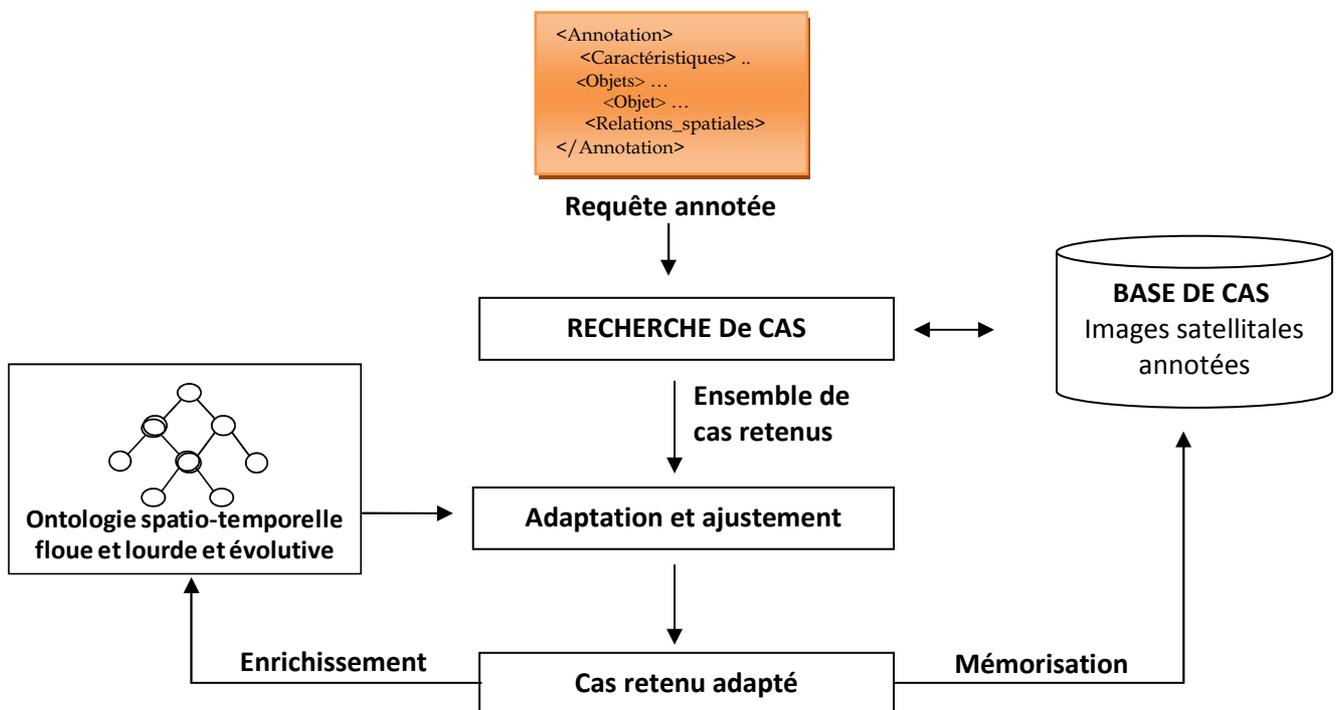


Figure 4.7 : Architecture du système d'interprétation d'images proposé

Le raisonnement à base de cas a pour objectif de trouver une solution, noté $Sol(cible)$ à un cas cible, noté *cible*. C'est un processus classiquement composé de trois opérations principales : la remémoration, l'adaptation et l'apprentissage.

La phase de remémoration permet de trouver les cas sources les plus similaires au cas cible, notés ($srce, Sol(srce)$).

La phase d'adaptation essaye d'adapter $Sol(srce)$ pour trouver une solution au problème *cible*. Un problème d'adaptation est alors un triplet ($srce, Sol(srce), cible$), et se résout en une solution $Sol(cible)$.

L'étape d'apprentissage consiste à mettre à jour les connaissances du système à l'issue du raisonnement. Elle est habituellement réalisée en mémorisant dans la base de cas le cas ($cible, Sol(cible)$).

Les phases de remémoration et d'adaptation se basent principalement sur les connaissances issues de l'ontologie de domaine.

4.2. Structure des cas

On distingue deux types de cas : le cas cible et le cas source.

Structure de cas cible : Un cas cible est représenté par l'annotation modélisant le contenu sémantique d'une image de télédétection, noté **Cible** ($A_c(\mathbf{Ob}_c, \mathbf{R}_c, C_c) ; I_c(\emptyset)$), avec :

- \mathbf{Ob}_c est l'ensemble d'objets de l'image ;
- \mathbf{R}_c est l'ensemble des relations spatiales entre les objets ;
- C_c est l'ensemble des caractéristiques et les contraintes de l'image ;
- I_c est l'interprétation de cas qui correspond à sa solution, initialement vide ;

Structure de cas source : Un cas source est un cas mémorisé dans la base de cas, et est représenté par l'annotation de l'image et l'ensemble des interprétations, noté **Source** ($A_s(\mathbf{Ob}_s, \mathbf{R}_s, C_s) ; I_s(\mathbf{S}_s)$). La figure 4.8 illustre la structure des cas cible et source.

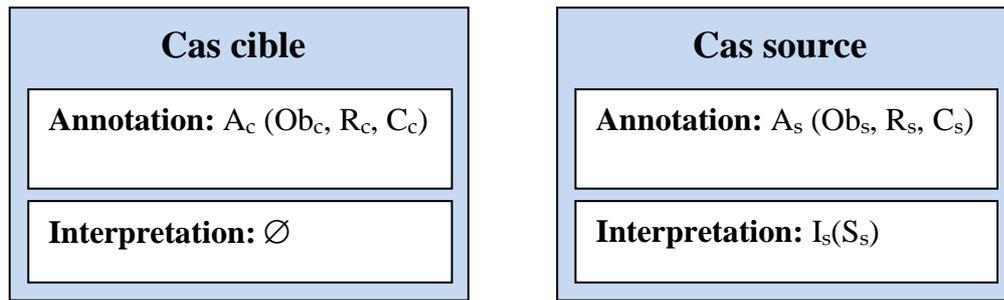


Figure 4.8 : Structure des cas cible et source

4.3. Remémoration guidée par ontologie

L'objectif de cette phase est de retrouver les cas similaires au cas cible à partir de la base de cas. Il s'agit de mesurer la similarité entre le cas cible et chaque cas mémorisé dans la base de cas.

Soit X un cas cible et Y un cas source :

$$X (A_X (Ob_X, R_X, C_X, Ct_X) ; I_X(\emptyset))$$

$$Y (A_Y (Ob_Y, R_Y, C_Y, Ct_Y) ; I_Y(S_Y))$$

Les deux cas X et Y sont similaires si leurs annotations A_X et A_Y sont similaires. Il s'agit de déterminer la similarité entre les objets, leurs relations et les contraintes des deux annotations.

La similarité globale entre les annotations de X et Y est :

$$SG(A_X, A_Y) = \omega_1 * SimO(Ob_X, Ob_Y) + \omega_2 * SimR(R_X, R_Y) + \omega_3 * SimC(C_X, C_Y) ; \sum_{i=1}^3 \omega_i = 1$$

Nous détaillons dans les sous-sections suivantes les différents types de similarités.

4.3.1. Similarité entre les objets

Il s'agit de mesurer la similarité entre les ensembles d'objets $Ob_X = \{o_{x1}, o_{x2}, \dots, o_{xn}\}$ et $Ob_Y = \{o_{y1}, o_{y2}, \dots, o_{ym}\}$. Ainsi, chaque objet de l'ensemble Ob_X doit être comparé avec tous les objets de l'ensemble Ob_Y .

Nous présentons tout d'abord la démarche que nous proposons pour mesurer la similarité $S(o_x, o_y)$ entre deux objets $o_x \in O_X$ et $o_y \in O_Y$:

Puisque o_x et o_y sont des instances de l'ontologie de domaine, nous proposons de les projeter sur l'ontologie et de déduire leur similarité comme suit :

- Si o_x et o_y appartiennent au même concept de l'ontologie alors : $S(o_x, o_y) = 1$;
- Si o_x et o_y appartiennent aux concepts différents de l'ontologie C_x et C_y respectivement, on applique alors la mesure de Wu et Palmer (1994) pour calculer la similarité structurelle $\text{Sim}_{\text{Struct}}(C_x, C_y)$ afin de déduire la proximité sémantique entre C_x et C_y . On ajoute à cette mesure une mesure de distance entre les attributs des objets o_x et o_y , notée $\text{Sim}_{\text{Att}}(o_x, o_y)$:

$$S(o_x, o_y) = \omega^* \text{Sim}_{\text{Struct}}(C_x, C_y) + \omega^* \text{Sim}_{\text{Att}}(o_x, o_y) \text{ avec } \sum \omega_i = 1,$$

avec,

$$\text{Sim}_{\text{Struct}}(C_x, C_y) = \frac{2 * \text{depth}(C)}{\text{depth}_c(C_x) + \text{depth}_c(C_y)} \text{ (Mesure de Wu et Palmer, 94), avec}$$

- C_x et C_y : les concepts de l'ontologie des objets O_x et O_y
- C est le Plus Petit Généralisant de C_x et C_y (en nombre d'arcs),
- $\text{depth}(C)$ est le nombre d'arcs qui sépare C de la racine
- $\text{depth}_c(C_i)$ est le nombre d'arcs qui séparent C_i de la racine en passant par C .

$$\text{Sim}_{\text{Att}}(o_x, o_y) = \frac{1}{n} \sum_{i=1}^n \omega * S(o_{x.a_i}, o_{y.a_i}), n \in \mathbb{N}^*, \sum \omega_i = 1 \text{ (Bergmann et al., 98)}$$

avec, $x.a_i$ est l'ensemble des attributs de l'objet x

$$S(o_{x.a_i}, o_{y.a_i}) = \frac{1}{1 + D(o_{x.a_i}, o_{y.a_i})} \text{ (Mesure de Johaniyak et al. 2005)}$$

$$D(o_{x.a_i}, o_{y.a_i}) = \sum_{i=1}^n |\mu_{o_x}(a_i) - \mu_{o_y}(a_i)| \text{ (Distance de Hamming)}$$

Concernant la similarité globale de deux ensembles d'objets $\text{Ob}_X = \{o_{x1}, o_{x2}, \dots, o_{xn}\}$ et $\text{Ob}_Y = \{o_{y1}, o_{y2}, \dots, o_{ym}\}$, nous adoptons la démarche suivante :

- 1) Pour chaque objet $o_{xi} \in O_X$ du cas X , on détermine l'objet $o_{yj} \in O_Y$ du cas Y le plus similaire comme suit :

$$\text{Sim}(o_{xi}, o_{yj}) = \max_{1 \leq k \leq n} (S(o_x, o_k)), o_k \in O_Y, n \in \mathbb{N}^*$$

- 2) On calcule alors la similarité globale entre les deux ensembles O_X et O_Y :

$$\text{Sim}_O(O_X, O_Y) = \sum_{i=1}^n \text{Sim}(o_{xi}, o_{yi}), o_x \in O_X, o_y \in O_Y, n \in \mathbb{N}^*$$

4.3.2. Similarité entre les relations spatiales

La mesure de similarité entre les objets est insuffisante pour déduire si deux cas sont similaires. La figure 4.9 présente un exemple illustratif de deux cas ayant les mêmes objets, mais leur répartition spatiale est différente. Cela montre que la prise en compte de similarité des relations spatiales entre les objets a une grande influence sur l'interprétation sémantique des cas. Nous proposons alors d'intégrer le contexte spatial des objets dans la mesure de similarité entre les cas.

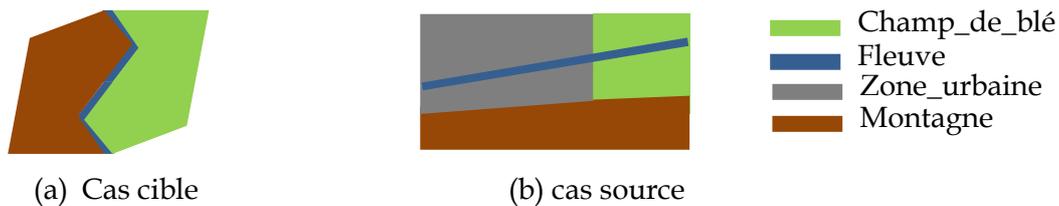


Figure 4.9. Importance des relations spatiales dans la mesure de relations spatiales

(a) Cas cible

- Objets {montagne ; fleuve ; champ_de_blé}
- Relations {Rc₁(Adjacent, montagne, fleuve) ; Rc₂(Adjacent, fleuve, champ_de_blé) ; Rc₃(Adjacent, champ_de_blé, montagne) ; Rc₄(A_Droite, fleuve, montagne) ; Rc₅(A_Gauche, fleuve, champ_de_blé) ; Rc₆(Disjoint, montagne, champ_de_blé) ; ...}

(b) Cas source

- Objets {montagne ; fleuve ; champ_de_blé ; zone_urbaine}
- Relations {Rs₁(Adjacent, montagne, champ_de_blé) ; Rs₂(Adjacent, zone_urbaine, champ_de_blé) ; Rs₃(Adjacent, zone_urbaine, montagne) ; Rs₄(bas, fleuve, zone_urbaine) ; Rs₅(bas, montagne, champ_de_blé) ; Rs₆(bas, montagne, champ_de_blé) ; ...}

Une relation spatiale $r(t, o_i, o_j)$ est définie par le type de relation t (adjacent, disjoint, à droite, à gauche, ...), et les objets o_i et o_j de la relation. Deux relations spatiales sont similaires si elles sont de même type et sont constituées des mêmes objets.

Soit $S_R(r_x(t_x, o_{xi}, o_{xj}), r_y(t_y, o_{yi}, o_{yj}))$ la similarité entre deux relations $r_x \in R_X$ et $r_y \in R_Y$ appartenant chacune aux cas X et Y respectivement. Pour chaque relation r_x du cas A, la relation correspondante dans le cas B est la relation r_y qui lui est similaire :

$$S_R(r_x(t_x, o_{xi}, o_{xj}), r_y(t_y, o_{yi}, o_{yj})) = \max_{1 \leq i \leq n} (S(r_x, r_i)), r_i \in R_y, n \in \mathbb{N}^*$$

$S(r_x, r_y)$ est déterminé selon les cas suivants:

(1) Même type de relation et mêmes objets :

$$(t_x=t_y), (o_{xi} = o_{yj}), (o_{xj} = o_{yj}) \rightarrow S(r_x, r_y) = 1$$

(2) Même type de relation et un seul couple d'objets différent :

$$(t_x=t_y), (o_{xi} \neq o_{yj}), (o_{xj} = o_{yj}) \rightarrow S(r_x, r_y) = 0.5$$

$$(t_x=t_y), (o_{xi} = o_{yj}), (o_{xj} \neq o_{yj}) \rightarrow S(r_x, r_y) = 0.5$$

(3) Même type de relation et tous les objets sont différents :

$$(t_x=t_y), (o_{xi} \neq o_{yj}), (o_{xj} \neq o_{yj}) \rightarrow S(r_x, r_y) = 0$$

(4) Différent type de relation et même objets : $(t_x \neq t_y), (o_{xi} = o_{yj}), (o_{xj} = o_{yj})$

Si t_x et $t_y \in \{\text{adjacent, proche, interconnecté...}\}$ ou $\{\hat{a} \text{ gauche, à droite}\}$
ou $\{\text{disjoint, loin}\} \rightarrow S(r_x, r_y) = 1$, sinon $S(r_x, r_y) = 0$

(5) Différent type de relation et un seul couple d'objets différent :

$$(t_x \neq t_y), (o_{xi} = o_{yj}), (o_{xj} = o_{yj}) \text{ ou } (t_x=t_y), (o_{xi} \neq o_{yj}), (o_{xj} = o_{yj})$$

Si t_x et $t_y \in \{\text{adjacent, proche, interconnecté...}\}$ ou $\{\hat{a} \text{ gauche, à droite}\}$
ou $\{\text{disjoint, loin}\} \rightarrow S(r_x, r_y) = 0.5$, sinon $S(r_x, r_y) = 0$

(6) Différent type de relation et tous les objets sont différents :

$$(t_x \neq t_y), (o_{xi} \neq o_{yj}), (o_{xj} \neq o_{yj}) \rightarrow S(r_x, r_y) = 0$$

Cette mesure permet de déterminer l'ensemble des relations similaires de deux cas, noté $Sr_i(r_x, r_y)$ avec i est le nombre de couples trouvés. La similarité globale des relations, notée $\text{SimR}_{(R_x, R_y)}$ est alors la somme des similarités $S_i(r_x, r_y)$:

$$\text{SimR}_{(R_x, R_y)} = \sum_{i=1}^n S_i(r_x, r_y), r_x \in R_x, r_y \in R_y, n \in \mathbb{N}^*$$

4.3.3. Similarité entre les contraintes

Soit $S_C(c_x, c_y)$ la similarité entre les contraintes et les caractéristiques des cas à aligner. Pour chaque contrainte $c_x \in C_X$ du cas A, la contrainte correspondant dans le cas B est la contrainte $c_y \in C_Y$ qui lui est la plus similaire :

$$Sc_{(C_x, C_y)} = \max_{1 \leq i \leq n} (Sc(c_x, c_i)), c_i \in C_y, n \in \mathbb{N}^*$$

$$Sc(c_x, c_y) = 1, \text{ si } c_x = c_y; 0 \text{ sinon}$$

La similarité globale de contraintes est alors :

$$\text{Sim}_{C_{(C_x, C_y)}} = \sum_{i=1}^n S_i(c, c_y), c_x \in C_X, c_y \in C_Y, n \in \mathbb{N}^*$$

4.4. La phase d'adaptation

4.4.1. Le problème d'adaptation

L'adaptation consiste à proposer une nouvelle solution $Sol(cible)$ au problème $cible$, à partir des solutions trouvées pour des problèmes proches, en modifiant la solution de $Sol(srce)$. Pour ce faire, l'adaptation consiste à effectuer un raisonnement par analogie et en s'appuyant sur des connaissances spécifiques au domaine. Le problème d'adaptation peut être lu sur le « carré d'analogie » illustré par la figure 4.10 (Chouraqui, 1986).

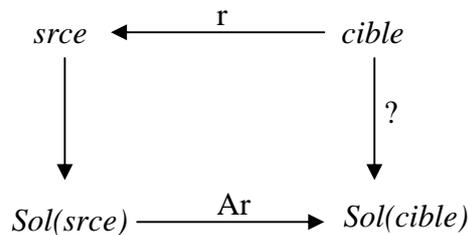


Figure 4.10 : Problème d'adaptation représenté par le carré d'analogie

Plusieurs approches d'adaptation ont été proposées dans la littérature en raisonnement à base de cas, à savoir (1) les approches transformationnelles et dérivationnelles (Carbonell, 1983) qui consistent à déterminer une modification ou une adaptation à appliquer à la solution du (ou des) cas source(s) remémoré(s), et (2) les approches par décomposition qui visent à décomposer la tâche d'adaptation en sous-tâches plus simples (Smyth et al., 2001 ; d'Aquin et al., 2005).

L'adaptation peut être formulée comme une tâche de recherche dans l'espace des solutions $Solutions$, où l'état initial est la solution $Sol(srce)$ d'un cas source remémoré et l'état final une solution $Sol(cible)$ pour le problème cible. Cette recherche s'effectue par l'application d'opérateurs d'adaptation, qui sont des transformations effectuées dans l'espace des solutions. Plusieurs types d'opérateurs d'adaptation sont utilisés dans la littérature pour modifier $Sol(srce)$:

- des opérateurs de copie, qui ne réalisent aucune transformation mais se contentent de recopier $Sol(srce)$,
- des opérateurs d'ajustement, qui modifient certaines valeurs de paramètres entrant en jeu dans $Sol(srce)$,
- des opérateurs de substitution, qui modifient $Sol(srce)$ en ajoutant, supprimant ou substituant certains de ses constituants,
- des opérateurs de transformation, qui opèrent une modification structurelle sur la solution $Sol(srce)$ (en changeant l'ordre des composants par exemple),
- des opérateurs de généralisation-spécialisation, qui exploitent une structuration hiérarchique de la base de cas pour généraliser tout ou partie de $Sol(srce)$ puis spécialiser en une solution pour le problème cible,

4.4.2. Technique d'adaptation adoptée

Nous proposons d'adopter l'approche d'adaptation par recherche en appliquant les opérateurs de reformulation, de substitution, de copie et selon le degré de similarité entre le cas source et le cas cible. Ces opérateurs permettent de transformer les différences dans les problèmes en différences dans les solutions à savoir la reformulation.

Une reformulation est un couple (r, Ar) où r est une relation entre problèmes et Ar est une fonction d'adaptation entre *solutions*. Une reformulation (r, Ar) signifie que : si un problème cible est lié à un autre problème $srce$ avec la relation r , on pourra obtenir $Sol(cible)$ à l'aide de la fonction Ar tel que : $Ar(srce, Sol(srce), cible) = Sol(cible)$.

Une relation r correspond à un ensemble de substitutions $\Sigma = \{\alpha \rightarrow \beta \mid \alpha, \beta \in V\}$. Une fonction d'adaptation Ar correspond également à un ensemble de substitutions $\Omega = \{\gamma \rightarrow \omega \mid \gamma, \omega \in V\}$. Les substitutions Ω sont générées automatiquement à partir de Σ :

$$\forall \alpha, \beta \in V, \alpha \rightarrow \beta \in \Sigma \rightarrow \beta \rightarrow \alpha \in \Omega$$

Les substitutions Σ peuvent être trouvées grâce aux relations entre les instances dans l'ontologie. En règle générale, une substitution $\beta \rightarrow \alpha$ est acceptée quand β et α sont similaires; c'est-à-dire que l'on pourrait tolérer le fait de remplacer α par β dans un cas source pour adapter sa solution au problème cible. On peut, par exemple, avoir des substitutions par généralisation (quand β subsume α) ou par spécialisation dans le cas inverse.

Conclusion

Dans ce chapitre, nous avons présenté une méthodologie de d'interprétation sémantique d'image de télédétection. Elle permet (1) d'extraire les informations et les connaissances de l'image satellitale, de (2) modéliser le contenu des images en se basant sur les connaissances de domaine, et en tenant compte du capteur utilisé pour l'acquisition et de la dimension temporelle, et enfin, (3) d'analyser le contenu de l'image et déduire si la scène qu'elle représente présente un risque naturel, en se basant sur un raisonnement inférentiel basé sur les connaissances du domaine et la base de cas.

Partie C

**Application
thématique**

Etude expérimentale de la méthodologie proposée : Etude de la susceptibilité de scène aux risques naturels ; Application à l'érosion

Il est bien connu que la dégradation et la disparition des couvertures végétales déclenchent et accélèrent les processus érosifs partout dans le monde, et plus particulièrement en Tunisie, sous l'action combinée de plusieurs facteurs à savoir le climat, la topographie, le couvert végétal et homme, entraînant de sérieux problèmes, tels que la réduction des terres arables, la diminution de la production agricole, la pollution des rivières, la sécheresse, etc.

Nous proposons alors d'appliquer notre méthodologie pour traiter ce problème afin de prévenir les zones à risque contre l'érosion en Tunisie et plus particulièrement à la région du Kef.

1. Les risques naturels

Un risque naturel est une menace qui découle de la conjonction d'un phénomène naturel (aléatoire) et de la présence de biens ou d'activités vulnérables.

D'après le dictionnaire de l'environnement, le risque naturel est la possibilité de survenance d'un évènement susceptible de porter atteinte à l'équilibre naturel. Le risque résulte de la conjonction d'un aléa et des enjeux en présence.

- L'aléa est la probabilité d'occurrence d'un phénomène donné. On parle toujours d'un aléa pour un phénomène et une durée donnés : aléa volcanique, aléa sismique, aléa mouvement de terrain, aléa inondation.
- Les enjeux sont les personnes, les biens, les équipements et l'environnement menacés par l'aléa, et susceptibles de subir des préjudices. Étroitement associée à la notion d'enjeux, la vulnérabilité mesure les conséquences de l'évènement sur les enjeux.

À titre d'exemple, un séisme de forte magnitude en plein désert causera peu ou pas de dommages, c'est pourquoi le risque pour ce phénomène est faible. À

contrario, le même séisme dans une zone urbanisée causera des dégâts conséquents, d'où un risque sismique élevé.

Les principaux risques naturels que nous pouvons distinguer sont : les inondations, les feux de forêt, les érosions des côtes, les érosions des sols, les mouvements de terrain, les tempêtes, les séismes et éruptions volcaniques, les sécheresses prolongées, etc. Certains de ces risques dépendent directement des activités des humains sur le globe, telles que les constructions des bâtiments, les habitudes agricoles et le surpâturage. Dans ce qui suit, nous nous intéressons aux risques d'érosions et à leurs facteurs déterminants.

1.1. Les risques d'érosion

L'érosion des sols par l'eau est un phénomène complexe qui résulte de divers processus (détachement, transport, dépôt) causés par l'action en général combinée de la pluie et du ruissellement, et dont l'expression varie en fonction de la résistance du milieu (sol nu, couvert végétal, techniques culturales) et de la topographie.

D'après le dictionnaire de l'environnement, l'érosion est l'action par laquelle divers éléments constituant les horizons superficiels de la couverture pédologique sont enlevés par le vent, la pluie, les rivières, les glaciers. Les facteurs principaux sont : la végétation, la couverture pédologique, la géomorphologie (pentes en particulier) et les impacts de l'utilisation des sols par l'homme.

En géomorphologie, l'érosion désigne un ensemble de processus physiques et chimiques de dégradation et de transformation du relief, et donc des roches, qui est causé par tout agent externe. Ce phénomène est considéré comme lent et progressif, mais ses conséquences peuvent être brutales et dangereuses.

Les facteurs essentiels qui conduisent à l'érosion sont le climat, la pente, la physique (dureté) et la chimie (solubilité par ex.) de la roche, l'absence ou non de couverture végétale et la nature des végétaux, l'action de l'homme (pratiques agricoles, urbanisation). Le rythme de l'érosion est conditionné par l'intensité de ses facteurs et peut, à long terme, araser des montagnes, creuser des vallées, faire reculer des falaises.

Les risques d'érosions se répartissent en deux classes essentielles : l'érosion des côtes et l'érosion des sols. L'érosion des sols se distribuent à leur tour en deux familles : l'érosion hydrique et l'érosion éolienne. Nous détaillons ces éléments dans les paragraphes qui suivent.

1.1.1. Erosion des côtes

Elle se manifeste par un ou des glissements ou des éboulements sur les berges d'un cours d'eau. Plusieurs facteurs naturels et anthropiques influencent la dynamique littorale. Ces facteurs agissent différemment selon les caractéristiques du

substrat et les caractéristiques du milieu. Tous les facteurs doivent être pris en considération lors de l'étude de l'érosion côtière.

Les caractéristiques du substrat :

Les principales caractéristiques du substrat qui contribuent à la sensibilité de l'érosion d'une portion côtière sont :

- La cohésion, qui est la capacité du matériel à se maintenir, à composer un ensemble, par exemple un cap rocheux comparé à une plage de sable;
- La texture du substrat, qui peut être grossière (galet et gravier), moyenne (sable) ou fine (limon, argile), et qui influence la capacité d'une vague à transporter le substrat;
- La structure du substrat joue aussi sur la sensibilité à l'érosion. Un substrat dont la structure est stratifiée répond différemment à l'érosion qu'un substrat dont la structure est massive;
- Les propriétés mécaniques sont des caractéristiques du substrat à considérer pour évaluer la sensibilité à l'érosion côtière. Par exemple, un substrat peut être imperméable ou à l'inverse poreux, ou encore il peut être consolidé ou à l'inverse plastique;
- D'autres caractéristiques comme l'uniformité du substrat, la nature, la composition et l'épaisseur des couches de sédiments sont déterminantes dans le comportement du substrat par rapport à l'érosion côtière.

Les caractéristiques du milieu :

La sensibilité d'un type de côte à l'érosion côtière dépend de sept principales caractéristiques propres au milieu.

- Le tracé du littoral décrit la linéarité de la rive. Par exemple, elle peut être rectiligne ou non, en retrait ou en saillie (cap ou péninsule), etc.;
- La topographie locale de l'arrière-plage, qui représente les terres émergées, influence la réaction du milieu à l'érosion. En effet, une terrasse (basse, moyenne ou haute), une plaine alluviale, etc., sont susceptibles de réagir de manière différente à l'érosion côtière;
- La topographie du fond marin adjacent au rivage modifie la dynamique littorale. Entre autres, l'inclinaison de la pente (forte ou faible, régulière ou irrégulière), l'étendue (largeur) de l'estran en rapport avec l'amplitude de la marée et la présence d'îles y contribuent;
- Le niveau d'exposition aux vents dominants et aux vents forts (de tempête), par la fréquence et l'intensité des vents, joue aussi sur la sensibilité du milieu à l'érosion côtière;
- La présence d'un cours d'eau peut soit amener des sédiments, soit favoriser l'érosion;
- Le drainage du milieu s'y ajoute, comme la qualité du drainage (des dépôts bien ou mal drainés), la profondeur de la nappe phréatique, la présence de tourbières, de ravins, etc.;

- L'intensité de l'occupation du territoire comprend le bâti et la pression que celui-ci exerce sur le milieu, la présence ou l'absence de végétation à l'état naturel (par opposition au gazon), les aires d'accès et les aménagements (passages répétés de piétons et de VTT sans rampe d'accès).

1.1.2. L'érosion hydrique

Le taux et l'amplitude de l'érosion hydrique dépendent des facteurs suivants:

L'intensité des précipitations et le ruissellement :

Il s'agit de l'impact des gouttes de pluie qui peut briser les agrégats et disperser les particules de sol. Les particules les plus fines, comme les sables fins, limons, argiles et la matière organique, peuvent facilement être emportées lors de l'éclaboussement des gouttes et par le ruissellement; des gouttes de pluie plus fortes (dissipant plus d'énergie) et un plus grand ruissellement seront nécessaires pour déplacer les plus grosses particules de sable et gravier.

Le gradient et la longueur de la pente :

Il est bien entendu que, plus la pente est dure, plus l'eau érodera le sol. L'érosion hydrique augmente aussi avec la longueur de la pente à cause de l'augmentation du ruissellement.

La Végétation :

Le risque d'érosion augmente lorsque le sol n'a qu'un faible couvert végétal ou de résidus. Les résidus et la végétation protègent le sol de l'impact des gouttes de pluie et de l'éclaboussement, tendent à ralentir la vitesse de l'eau de ruissellement et permettent une meilleure infiltration.

L'efficacité du couvert de végétation et de résidus à réduire l'érosion dépend du type, de l'étendue et de la densité du couvert végétal. La végétation et les résidus combinés, couvrant complètement le sol, interceptent la pluie et sont le moyen le plus efficace pour réduire les pertes de sol.

La capacité de maîtriser l'érosion de toute culture dépend aussi de la protection disponible à diverses périodes de l'année en relation avec la quantité de précipitation à caractère érosif durant ces périodes. A cet égard, les plantes qui assurent un bon couvert végétal durant la majeure partie de l'année, maîtrisent mieux l'érosion que les cultures qui laissent le sol nu durant une longue période, particulièrement pendant les périodes de précipitations très érosives.

1.1.3. L'érosion éolienne

Le taux et l'amplitude de l'érosion éolienne sont contrôlés par les facteurs suivants:

Susceptibilité du sol à l'érosion :

Le vent peut mettre en suspension de très fines particules et les transporter sur de grandes distances. Les particules moyennes et fines peuvent être soulevées et redéposées alors que les grosses particules peuvent être soufflées en surface (saltation). L'abrasion qui en résulte peut réduire la dimension des particules de sol et augmenter d'autant sa susceptibilité à l'érosion.

La rugosité de la surface :

Les sols à la surface non rugueuse offrent très peu de résistance au vent. Toutefois, avec le temps, les rugosités peuvent se combler et s'aplanir par abrasion. Ceci résulte en une surface lisse, plus susceptible au vent. Un travail excessif du sol peut contribuer à briser la structure du sol et ainsi augmenter l'érosion.

Le climat :

La vitesse et la durée du vent ont un effet direct sur l'érosion du sol. Le taux d'humidité peut être très faible à la surface d'un sol excessivement drainé ou durant une sécheresse. Les agrégats s'émiettent en particules qui peuvent être transportées par le vent.

La longueur exposée des champs :

Le manque de brise-vent (arbres, arbustes, résidus, etc.) permet au vent de mettre en mouvement des particules de sol sur de grandes distances, augmentant ainsi l'abrasion et l'érosion du sol. Les monticules sont habituellement les plus exposés et en souffrent davantage.

Le couvert végétal :

Le manque de couvert végétal permanent à certains endroits a conduit à une érosion éolienne considérable. Les sols nus, secs et exposés sont les plus susceptibles ; toutefois, les cultures produisant peu de résidus offrent parfois une résistance insuffisante. De même les cultures qui produisent beaucoup de résidus ne procurent pas une protection adéquate dans les cas extrêmes.

Le couvert végétal le plus efficace devrait comprendre un réseau de brise-vents végétaux couplé à un bon travail du sol, une bonne régie des résidus et un choix approprié des cultures.

2. Application de la méthodologie

Nous proposons d'appliquer notre méthodologie pour étudier les zones à risques d'érosion pour la région de Kef située au nord-ouest de la Tunisie. En effet, nous avons enrichi l'ontologie par des connaissances de domaine de l'érosion. Puis, nous avons appliqué la méthodologie d'interprétation selon deux modes : le mode *offline* et le mode *online*. Nous détaillons chacune de ces étapes dans les sous-sections suivantes.

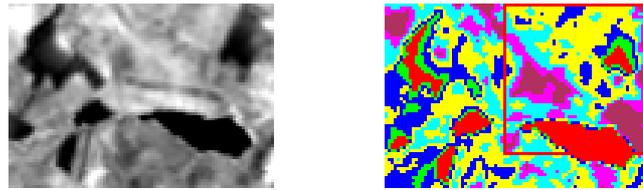
2.1. Le mode *offline*

Il s'agit d'un processus d'apprentissage permettant d'initialiser notre méthodologie d'interprétation des images de télédétection. Pour ce faire, nous avons construit une base de cas modélisant des scènes contenant des zones à risques d'érosion. En effet, la phase d'annotation sémantique est appliquée sur ces images pour (1) générer leurs annotations correspondantes, et (2) alimenter la base de cas. A ce stade, l'interprétation des annotations s'effectue à travers le raisonnement inférentiel des connaissances issues de l'ontologie de domaine. Nous avons établi une base de 30 cas traitant différents types de cas d'érosions.

2.1.1. La phase d'annotation

Nous présentons un exemple illustrant le processus d'annotation de l'image de télédétection qui se déroule en ces étapes :

- 1) Soit l'image de télédétection représentant une scène par un capteur SPOT en mois de mars (Figure 5.1(a)).
- 2) Extraction des régions et leurs primitives : Segmentation supervisée de l'image avec l'algorithme ISODATA, et extraction des primitives avec l'outil ENVI (Figure 5.1(b)).



(a)

(b)

Figure 5.1 : (a) Image de télédétection issues de capteur Spot en mois de mars, (b) régions extraites de l'image.

- 3) Modélisation floue des régions: Cette étape permet d'établir une représentation gaussienne des primitives de régions extraites. Le tableau 5.1 résume les caractéristiques calculées de deux régions.

Region A			
Primitives	Valeur		
Radiométrie	$\mu_{R_S_B1_P}$ 	$\mu_{R_S_B2_P}$ 	$\mu_{R_S_B3_P}$
NDVI	$\mu_{NDVI_S_B1_P}$ 	$\mu_{NDVI_S_B2_P}$ 	$\mu_{NDVI_S_B3_P}$
Texture	$\mu_{Tex_B1_P}$ 	$\mu_{Tex_S_B2_P}$ 	$\mu_{Tex_S_B3_P}$

Tableau 5.1 : Primitives floues d'une région

- 4) Matching flou :

Cette étape permet de déterminer le concept (de l'ontologie) approprié pour chaque région de l'image. Nous mesurons la similarité floue entre la région requête et les concepts de l'ontologie. Nous retenons le concept ayant le degré d'appartenance le plus élevé. Ce processus est appliqué à tous les régions extraites afin de déterminer leurs classes correspondantes de l'ontologie. Le tableau 5.2 résume les similarités effectués pour la région A. La similarité globale entre une région et un concept est la somme de similarité entre les primitives de la région (PR) et les propriétés du concept (PC) :

$$S(R, C_i) = \frac{\sum_{i=1}^{np} S(PR_i, PC_i)}{np}$$

Région	Matching flou	Concept retenu
A 	S(A, Lac)=0.8314 S(A, Forêt)=0.3279 S(A, Sol nu) = 0.01958 ...	Lac, Score = 0.8314
B 	S(B, Champ_de_blé)=0.74565 S(B, forêt)=0.4987 ...	Champ de blé, Score = 0.74565
C 	S(B, Sol nu)=0.6980 S(B, forêt)=0.2153 ...	Sol nu, 0.6980

Tableau 5.2 : Similarités floues entre les régions et les concepts ontologiques

5) Annotation résultante

C'est le résultat du processus du matching flou. Nous obtenons alors une méta-information contenant les régions de l'image, leurs primitives et les relations spatiales entre eux. La figure 5.2 présente un extrait de l'annotation résultante au format XML.

```

<Annotation>
  <Caractéristiques>
    <Capteur> SPOT </Capteur>
    <nb_Bandes> 3 </nb_Bandes>
    <Résolution> 20 mètres </Résolution>
    <Géo-Ref>
      <Proj> UTM, Zonz 32N </Proj>
      <Datum> Carthage </Datum>
    </Géo-Ref>
  </Caractéristiques>
  <Objets>
    <Objet>
      <id_objet>A</id_objet>
      <concept_ontologie>
        <Nom_concept> Lac </Nom_concept>
        <Degré_Appartenance> 0.8314 </Degré_Appartenance>
        <Primitives>
          <Radiométrie>
            <μR,S_B1,P>
              <x> 50.02905419537401, 52.684195066437006,
                53.339335937499996, 54.994476808562986, 55.64961767962598,
                56.30475855068897, 56.95989942175196</x>
              <y> 0,563293138186075      0,666493599425726
                0,768748114039742      0,963610881410243
                0,987200780122373 0,992602841981015
                0,973078629388837      0,942095985691319
                0,900776753677940      0,897942942461284 </y>
            </μR,S_B1,P>
            <μR,S_B2,P> ... </μR,S_B2,P>
          </Radiométrie>
        </Primitives>
      </concept_ontologie>
    </Objet>
  </Objets>
</Annotation>

```

```

...
</Radiométrie>
<Texture> ... </Texture>
<Forme> ... </Forme>
</Primitives>
</Objet>
<Objet>
  <id_objet>B</id_objet>
  <concept_ontologie>
    <Nom_concept> Champ de blé </Nom_concept>
    <Degré_Appartenance> 0.74565 </Degré_Appartenance>
  <Primitives>
    <Radiométrie>
      <μR,S_B1_P>
        <x> 20.02905419537401, 22.684195066437006,
        23.339335937499996, 24.994476808562986, 25.64961767962598,
        26.30475855068897, 26.95989942175196</x>
        <y> 0,563293138186075      0,666493599425726
              0,768748114039742    0,963610881410243
              0,987200780122373    0,992602841981015
              0,973078629388837    0,942095985691319
              0,900776753677940    0,897942942461284 </y>
      </μR,S_B1_P>
      <μR,S_B2_P> ... </μR,S_B2_P>
    ...
  </Radiométrie>
  <Texture> ... </Texture>
  <Forme> ... </Forme>
</Primitives>
</Objet>
<Objet> ... <Objet>
</Objets>
<Relations_spatiales>
  <Relation_spatiale>
    <type>Adjacent </type>
    <Objet1> A </Objet1>
    <Objet2> B </Objet2>
    <Degré_Appartenance> 0.7856</Degré_Appartenance>
  </Relation_spatiale>
  <Relation_spatiale>
    <type>Disjoint </type>
    <Objet1> A </Objet1>
    <Objet2> C </Objet2>
    <Degré_Appartenance> 0.9856</Degré_Appartenance>
  </Relation_spatiale>
  ...
</Relations_spatiales>
</Annotation>

```

Figure 5.2 : Extrait de l'annotation de l'image

L'annotation obtenue modélise les classes de l'image (région requête associée au concept correspondant de l'ontologie), les relations spatiales entre les classes, et les informations générales de la requête. Cette annotation constitue le « cas » sur laquelle se base le processus de raisonnement à base de cas et d'ontologie.

2.1.2. *La phase de raisonnement*

Elle se base sur les règles d'inférences issues de l'ontologie de domaine. Concernant l'exemple illustratif, la déduction suivante est ajoutée au cas :

Risque_Erosion = fort, Objet = B

Cette affirmation est déduite de l'application de la règle de l'ontologie : *Si un champ de blé est situé à côté d'une rivière et le type de sol est argileux, alors le risque d'érosion est élevé.*

La phase d'apprentissage a permis de construire une base de 30 cas modélisant différents cas d'érosion.

2.2. *Le mode online*

Le mode *online* a pour objectif de déduire si une requête (image de télédétection) soit susceptible à un risque d'érosion en se basant sur le raisonnement à base cas. Ce mode se déroule en quatre étapes :

- 1) Annotation de la requête : construction d'un nouveau cas Cas_{Req} .
- 2) Recherche des cas similaires au cas Cas_{Req} .
- 3) Adaptation de l'interprétation au Cas_{Req} .
- 4) Mise à jour de la base de cas.

Illustration :

Soit l'image requête suivante modélisant une scène qu'on veut savoir la présence de risque d'érosion. La scène est un champ de blé près d'une montagne (Figure 5.3).

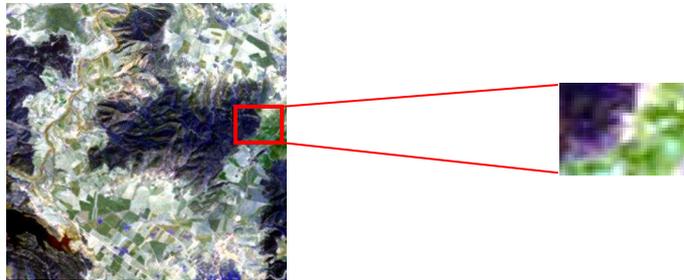


Figure 5.3 : Image requête

Informations liées à l'image :

- Capteur : SPOT, 3 bandes
- Projection : UTM, Zone 32 North
- Résolution : 20 Meters
- Datum : Carthage
- Pente montagne : 15°

La figure 5.4 présente un résumé du cas établi.

<p>(C_{im}) Acquisition et caractéristiques</p> <ul style="list-style-type: none">• Date d'acquisition : 2005• Résolution : 20 mètres• Géo-référencement : UTM, Zone 32 North• Capteur : Spot• Nombre de bandes : 3
<p>(C_{im}) Ensemble d'objets de l'image, et leurs primitives</p> <p>Objet 1 (A) : Montagne { ... }</p> <p>Objet 2 (B) : Champs de blé { ... }</p>
<p>(R_{im}) Relations spatiales</p> <p>Proche (A, B)</p>
<p>Interprétation</p> <p>∅</p>

Figure 5.4 : Cas représentant l'image requête

Le tableau 5.3 présente des similarités calculées entre le cas cible et les cas sources (dans base de cas) :

Cas 1 : 0.001	Cas 16 : 1.004
Cas 2 : 0.240	Cas 17 : 1.000
Cas 3 : 0.038	Cas 18 : 0.001
Cas 4 : 0.409	Cas 19 : 0.001
Cas 5 : 2.789	Cas 20 : 0.001
Cas 6 : 3.309	Cas 21 : 0.300
Cas 7 : 1.98	Cas 22 : 0.003
Cas 8 : 7.607	Cas 23 : 0.006
Cas 9 : 1.119	Cas 24 : 0.009
Cas 10 : 0.040	Cas 25 : 2.000
Cas 11 : 0.078	Cas 26 : 1.000
Cas 12 : 1.098	Cas 27 : 1.000
Cas 13 : 0.040	Cas 28 : 3.000
Cas 14 : 0.040	Cas 29 : 1.000
Cas 15 : 0.040	Cas 30 : 0.008

Tableau 5.3 : Récapitulatif des similarités entre le cas cible et la base de cas

Nous avons fixé un seuil de sélection des cas candidats supérieurs ou égaux à 30%. D'après ces résultats, le cas 8 est le plus similaire au cas cible. Nous retenons alors ce cas et nous adaptons son interprétation au cas cible. La figure 5.5 présente le cas cible après la phase d'adaptation :

<p>(C_{im}) Acquisition et caractéristiques</p> <ul style="list-style-type: none"> • Date d'acquisition : 2005 • Résolution : 20 mètres • Géo-référencement : UTM, Zone 32 North • Capteur : Spot • Nombre de bandes : 3
<p>(C_{im}) Ensemble d'objets de l'image, et leurs primitives</p> <p>Objet 1 (A) : Montagne {...}</p> <p>Objet 2 (B) : Champs de blé {...}</p>
<p>(R_{im}) Relations spatiales</p> <p>Proche (A, B)</p>
<p>Interprétation</p> <p>Rique_erosion : Moyen, Région B</p>

Figure 5.5 : Cas cible après mise à jour

Ensuite, nous ajoutons ce nouveau cas à la base de cas. Nous avons essayé de simuler 50 nouveaux cas, en effet, nous avons repéré plusieurs zones de la région d'étude à risque moyen et fort d'érosion.

Conclusion

Dans ce chapitre, nous avons essayé d'appliquer notre méthodologie à l'étude de risque d'érosion dans la région de kef de la Tunisie. En fait, elle a permis de générer des annotations sémantiques aux scènes analysées et de déterminer des zones à risques d'érosion afin de les surveiller au plus tôt possible.

Conclusions et perspectives

1. Rappel des objectifs

Le cadre général de notre thèse est l'interprétation sémantique d'images de télédétection guidée par ontologie. Les objectifs de la thèse sont les suivants :

Cadre formel : Proposition d'une ontologie modélisant le contenu de l'image de télédétection, et prenant en compte de ses spécificités (capteurs, temps, imperfections et contraintes) et les connaissances de domaine.

Cadre fonctionnel : Proposition d'une méthodologie permettant de générer des annotations modélisant le contenu de l'image satellitale en se guidant par des l'ontologie proposée, et permettant, à partir d'un raisonnement inférentiel sur ces connaissances, de déduire la susceptibilité d'une scène satellitale aux risques naturels.

Cadre applicatif : Application de la méthodologie proposée pour étudier la susceptibilité d'une scène au risque d'érosion. Etude de cas de la région de Kef-Tunisie.

2. Contributions

Dans ce travail de thèse, nous avons proposé une ontologie spatio-temporelle, floue, lourde et évolutive représentant le contenu sémantique des images de télédétection.

Nous avons proposé un processus de développement basé sur la réutilisation des ressources ontologiques existantes, l'apprentissage à partir des images satellitales (de référence), et la formalisation des connaissances *a priori* de domaine (risques naturels, connaissance sur les scènes capteurs d'acquisition, etc.). Ce processus est comme suit :

- 1) Choix de l'ontologie de noyau
- 2) Evolution structurelle de l'ontologie de noyau
 - a. Modélisation floue.
 - b. Ajout de la dimension temporelle et spectrale
- 3) Enrichissement conceptuel, intensionnel et relationnel de l'ontologie de noyau
 - a. Découverte des ressources ontologiques.
 - b. Evaluation : détermination des parties à réutiliser de chaque ressource en utilisant des techniques d'alignement d'ontologies.
 - c. Intégration et placement de nouveaux concepts.
- 4) Enrichissement axiomatique

Nous avons validé ce processus avec une ontologie de noyau tirée du travail de (Durand et al., 2007) dans le cadre du projet Fodumust, des ressources ontologiques dans le domaine géographique, et des connaissances géographique sur le domaine de risque naturel d'érosion. Nous avons obtenu des résultats satisfaisants pour un travail initial avec l'ontologie.

Comme cadre méthodologique de travail, nous avons considéré cette ontologie comme support de connaissances pour (1) guider l'annotation sémantique des images de télédétection, et (2) guider le raisonnement inférentiel renforcé par le raisonnement à base de cas afin d'étudier la susceptibilité de scène aux risques naturels.

Nous avons alors proposé un processus d'annotation permettant de générer une méta-information modélisant le contenu sémantique d'une image. Ce processus est basé sur la similarité floue entre les régions d'une image requête avec les concepts de l'ontologie de domaine proposée. Ce processus est illustré comme suit :

- 1) Extraction des informations de l'image de télédétection :
 - a. Extraction des régions de l'image ;
 - b. Extraction des primitives visuelles de chaque région ;
 - c. Extraction des relations spatiales entre les régions ;
- 2) Modélisation floue des régions extraites.

- 3) Matching floue : Détermination du concept (de l'ontologie) approprié pour chaque région de l'image.
- 4) Etablissement de l'annotation sémantique.

L'annotation obtenue modélise les classes de l'image (région requête associé au concept correspondant de l'ontologie), les relations spatiales entre les classes, et les informations générales de l'image. Cette annotation constitue le « cas » la sur la quelle se base le processus de raisonnement à base de cas et d'ontologie.

3. Perspectives

Nous proposons d'étendre notre travail en ajoutant l'aspect dynamique à notre ontologie proposée, pour qu'elle soit capable de détecter les changements d'une série temporelle de scène. Cela nécessite d'étudier la possibilité d'intégrer cet aspect dans la définition conceptuelle, intentionnelle et relationnelle de l'ontologie.

Dans un cadre applicatif, nous projetons d'examiner d'autres cas de risques naturels à savoir les inondations, les feux de forêts, la désertification, etc. Cela nécessite d'enrichir l'ontologie par de nouvelles ressources et connaissances de chaque domaine, et bien évidemment, d'adapter les mécanismes de raisonnement selon les connaissances de domaines.

Annexes

Annexe 1 : Les techniques d'alignement

Les techniques d'alignement appelées méthodes locales, rassemblent plusieurs méthodes : les méthodes terminologiques, les méthodes structurelles, les méthodes extensionnelles et les méthodes sémantiques (Euzenat, 2007).

1. Méthodes terminologiques

Ces méthodes comparent les chaînes de caractères (exemple les labels qui décrivent les concepts) afin d'en déduire la similarité (ou dissimilarité) en exploitant les relations d'hyponymie¹ ou d'hyperonymie². Certaines méthodes terminologiques se basent sur la comparaison des chaînes de caractères et sont appelées les méthodes syntaxiques. Pour ces méthodes deux chaînes qui partagent des caractères ou des mots en commun seront considérées comme similaires. D'autres méthodes terminologiques, font appel à une base de données lexicale sous forme de réseau sémantique. Ce sont les méthodes linguistiques. Ces méthodes calculent la similarité entre deux chaînes en fonction des relations qu'elles retiennent.

1.1 Méthodes syntaxiques

Les méthodes syntaxiques se basent sur la comparaison des entités à aligner par le calcul de la valeur de la similarité des entités. Ces entités sont caractérisées par des noms, des étiquettes, de commentaires, des descriptions, etc. Ces méthodes se déclinent en deux sous-catégories (SHVAIKO, 2005):

- La première sous-catégorie englobe des méthodes qui comparent des termes en se basant sur les caractères contenus dans ces termes.
- La deuxième sous-catégorie : utilise les distances basées sur les tokens.

Les méthodes syntaxiques, analysent la structure de ces chaînes afin de calculer des mesures de similarité. Ces méthodes profitent de l'ordre des caractères dans la chaîne, le nombre d'apparitions d'une lettre dans une chaîne. En revanche, elles n'exploitent pas la signification des termes. On trouve plusieurs mesures garantissant le calcul de la valeur de similarité ou de dissimilarité entre deux chaînes de caractères dans la littérature. Parmi les

¹ Relation sémantique hiérarchique entre les mots : le sens du premier englobe le second

² Le sens du premier est inclus dans le sens du second

mesures les plus utilisées dans le cadre de l'alignement d'ontologies, on cite la distance de HAMMING, la similarité de JACCARD, la distance d'édition, la distance de LEVENSHTTEIN, etc. Ces mesures sont détaillées dans ce qui suit :

Dans la suite, une chaîne de caractères est considérée comme un ensemble de caractères, noté S.

- **La distance de HAMMING** : est une mesure de dissimilarité qui comptabilise le nombre de positions dans lesquelles les deux chaînes de caractères sont différentes [Hamming,1950].

$$\overline{\delta}_{HAMMING} : S, S \rightarrow [0,1] /$$

$$\overline{\delta}_{HAMMING}(s, s') = \frac{(\sum_{i=1}^{\min(|s|, |s'|)} s[i] \neq s'[i]) + ||s| - |s'|}{\max(|s|, |s'|)}$$

- **Similarité de JACCARD** : Étant donné deux chaînes de caractères s et s'. Les deux ensembles S et S' représentent, respectivement, les caractères de s et s'. La similarité de JACCARD est définie comme suit:

$$\overline{\sigma}_{Jaccard}(s, s') : S \times S \rightarrow [0,1] \text{ tel que } \overline{\sigma}_{Jaccard}(s, s') = \frac{|S \cap S'|}{|S \cup S'|}$$

- **La distance d'édition** : Etant donné un ensemble Op d'opérations d'édition $op : S \rightarrow S$ et une fonction de coût $w : op \rightarrow \mathbb{R}$ tel que pour chaque paire de chaîne de caractères il existe une séquence d'opération d'édition permettant la transformation de la première chaîne en la seconde (et vice versa). La distance d'édition est

$$\overline{\delta}_{de} : S \times S \rightarrow [0,1], \text{ avec } \overline{\delta}_{de}(s, s')$$

est le coût minimal de la séquence d'opérations d'édition transformant s en s' /

$$\overline{\delta}_{de}(s, s') = \min op_i; op_n (...op_1(s)) = s' (\sum_{i \in I} w op_i)$$

- **La distance LEVENSHTTEIN** : c'est une mesure avec un nombre minimal d'opérations d'insertions, de suppressions et de substitutions de caractères nécessaires pour la transformation d'une chaîne à une autre [Levenshtein,1966].

- **La mesure de Jaro** : c'est une mesure définie par la fonction $\overline{\sigma}_{Jaro} : S \times S /$

$$\overline{\sigma}_{Jaro}(s, s') = \frac{1}{3} \times \left(\frac{|com(s, s')|}{|s|} + \frac{|com(s, s')|}{|s'|} + \frac{|com(s, s')| - |transp(s, s')|}{|com(s, s')|} \right)$$

avec $s[i] \in com(s, s')$ si et seulement si $\exists j \in [i - \min(|s|, |s'|) / 2, i + \min(|s|, |s'|) / 2]$ et $transp(s, s')$ sont les éléments de $com(s, s')$ qui apparaissent dans des ordres différents

dans s et s'.

- **La distance de Jaro-Winkler** : est une amélioration de la mesure de Jaro, qui favorise les correspondances entre les chaînes avec des préfixes communs [Winkler,1999]

$$\overline{\sigma}_{Jaro-Winkler} : S \times S$$

$$\overline{\sigma}_{Jaro-Winkler}(s, s') = \overline{\sigma}_{Jaro(s, s') + P \times Q} \times \frac{(1 - \overline{\sigma}_{Jaro(s, s')})}{10}$$

avec P la longueur du préfixe commun et Q une constante.

Les mesures citées auparavant sont les plus utilisées lors de la comparaison de deux chaînes de caractères courtes. Dans le cas de texte longs ou de documents textuels, ces chaînes sont découpées en plusieurs morceaux appelés tokens, et la similarité est mesurée grâce aux mesures de similarité basées sur les tokens.

- **La mesure de Q-Gram:**

$$\sigma_{Q-Gram} : S \times S$$

$$\sigma_{Q-Gram}(s, s') = \frac{|qgram(s, n) \cap qgram(s', n)|}{\min(|s|, |s'|) - n + 1}$$

où n est la longueur d'un qgram.

Le calcul de la similarité entre les commentaires des entités à aligner est accompli par la mesure Q-GRAM [Ukkonen,1992]. Elle est exploitée pour la comparaison des chaînes de caractères composées de plusieurs mots. Cette mesure calcule le nombre de qgrams en commun, c-à-d., les séquences de n caractères figurant à la fois dans deux chaînes de caractères.

- **Distance de MONGE-ELKAN** : La distance de MONGE-ELKAN estime la similarité entre deux chaînes s et s', représentées par leurs sous chaînes respectives s = ai... ak et s' = bi...bt. La mesure de MONGE-ELKAN

$$\delta_{MONGE-ELKAN} : S \times S \text{ tel que}$$

$$\delta_{MONGE-ELKAN}(s, s') = \frac{1}{|s|} \left(\sum_{i=1}^{|s|} \max_{j=1}^{|s'|} sim(a_i, b_j) \right)$$

La sim(ai,bj) est la similarité entre les sous chaînes ai et bj. Cette similarité est calculée par l'intermédiaire d'une des mesures de similarité, telles que JACCARD, JARO, JARO-WINKLER, etc.

La distance de MONGE-ELKAN est une approche hybride qui permet de calculer la similarité entre deux chaînes de manière récursive [Monge,1996a], en analysant les sous-chaînes des chaînes initiales.

1.2 Méthodes linguistiques

La similarité entre deux entités représentées par des termes (ou mots) peut aussi être calculée en les analysant à l'aide des méthodes linguistiques qui prennent en charge les propriétés expressives et productives du langage naturel et peuvent être intrinsèques ou extrinsèques (Euzenat, 2007).

- **Les informations intrinsèques** : ce sont les propriétés linguistiques internes des termes, telles que des propriétés morphologiques ou syntaxiques. Ces méthodes examinent la forme conforme et représentative d'un mot ou d'un terme (lemme) en exploitant ses variantes linguistiques (lexème). La similarité entre deux termes est mesurée en comparant leurs lemmes.
 - La troncature : c'est une approche automatique qui permet de déterminer la forme radicale d'un terme. Cette forme est retenue à partir d'une forme infléchie ou dérivée d'un mot donné. Les radicaux trouvés n'ont pas besoin d'être semblables à la racine morphologique du mot. Il suffit que les mots similaires soient adjoints, à un même radical, même si, ce radical n'est pas une racine de mot valide sémantiquement.
 - La lemmatisation : c'est une approche qui détermine le radical exact d'un mot, par l'extraction de la catégorie lexicologue d'un mot, appliquant ainsi des règles de normalisation différentes pour chaque partie du discours. Cette approche nécessite la connaissance de la grammaire ainsi que les différentes règles d'un langage.
- **Les informations extrinsèques** : Ces méthodes calculent la valeur de similarité entre deux termes en employant des ressources externes telles que les dictionnaires, les lexiques ou les vocabulaires. La similarité entre deux termes est calculée en exploitant les liens sémantiques existants dans ces ressources externes. Ces liens regroupent les synonymes pour l'équivalence, les liens d'hyponymes ou d'hyperonymes pour la subsomption. Parmi les ressources externes employées, nous trouvons :
 - l'API WORDNET (Miller ,1995) est un système lexicologique qui a été utilisé pour chercher des relations telles que la synonymie entre les termes.
 - WORDNET un système exploité pour calculer la distance sémantique entre les termes.

Les ressources externes utilisées dans les méthodes extrinsèques peuvent aussi être des vocabulaires ou des dictionnaires multi-lingues, ou d'autres systèmes tels que EuroWORDNET et POLYLEX.

2. Méthodes structurelles

Les méthodes structurelles déterminent la similarité entre deux entités en fonction des informations structurelles. En effet, les entités sont reliées entre elles par des liens sémantiques ou syntaxiques. Ces liens établissent ainsi une hiérarchie ou un graphe d'entités. Ces méthodes se subdivisent en deux familles : les méthodes structurelles internes et les méthodes structurelles externes. La première famille exploite uniquement les informations décrivant les attributs d'entités, tandis que la seconde famille considère des relations entre ces entités.

2.1 Méthodes structurelles internes

Les méthodes structurelles internes appelées aussi les approches basées sur les contraintes [Rahm , 2001], utilisent les informations communes dans les structures internes des entités pour le calcul de la similarité. Ces informations regroupent les attributs des entités qui sont les informations du co-domaine, de la cardinalité des attributs, des caractéristiques des attributs (la transitivité, la symétrie, etc.), et éventuellement des restrictions sur les attributs.

[Valtchev ,1999] a introduit la similarité entre deux types de données. Cette mesure est le rapport entre la différence des tailles des types et la taille de leur généralisation commune, sachant que la taille d'un type est définie comme la cardinalité de l'ensemble de valeurs qu'il procure et la taille de la généralisation commune dépend des types ces éléments.

2.2 Méthodes structurelles externes

Les méthodes structurelles externes exploitent les relations existantes entre les entités elles-mêmes. Ces relations contiennent des relations de subsomption (spécialisation ou Is a) ou de méréologie (part-whole). En plus de ces relations, les entités sont observées dans des structures hiérarchiques. La similarité entre les entités est déterminée en fonction de leurs positions dans leurs hiérarchies.

Pour [Euzenat,2007] deux entités peuvent être observées similaires dans les cas où :

- Leurs super-entités directes (ou leurs super-entités) sont similaires ;
- Leurs sœurs (ou toutes leurs sœurs, qui sont les entités ayant la même super-entité directe avec les entités en question) sont déjà similaires ;
- Leurs sous-entités directes (ou toutes leurs sous-entités) sont déjà similaires
- Leurs descendants (entités dans le sous-arbre ayant pour racine l'entité en question) sont déjà similaires ;
- Toutes leurs feuilles (les entités de même type qui n'ont aucune sous-entité, dans le sous-arbre ayant pour racine l'entité en question) sont déjà similaires.

- Toutes les entités dans les chemins de la racine aux entités en question sont déjà similaires.

Des combinaisons de ces différents cas sont aussi possibles. Les approches externes rencontrent des difficultés dans le cas où les hiérarchies à aligner possèdent des niveaux différents de granularité.

3. Méthodes extensionnelles

Les méthodes extensionnelles calculent la similarité entre deux entités en fonction de leurs extensions [Euzenat,2007] qui représentent les ensembles des instances de deux entités. Dans le cas où l'intersection des ensembles des instances est grande, les mesures utilisant des opérations sur les ensembles peuvent être exploitées.

Ces mesures comprennent la distance de HAMMING et la distance de JACCARD qui peuvent être adaptées pour développer des mesures extensionnelles, produisant la similarité de deux entités en fonction de la similarité entre les deux ensembles en se basant sur la comparaison exacte des éléments dans les deux ensembles. Dans le cas où les ensembles des instances n'ont aucune partie commune, ces mesures ne sont plus applicables. En 1999 Valtchev, a recommandé la similarité basée sur des correspondances (match-based similarity). La similarité entre deux ensembles est la similarité moyenne des éléments dans un ensemble de correspondance, noté PARING.

Similarité basée sur les correspondances : étant donné une fonction de similarité $\sigma : E \times E \rightarrow \mathfrak{R}$, la similarité de correspondances entre deux sous-ensembles E, où E est un ensemble d'entités, est une fonction de similarité

$$MSim : 2E \times 2E \rightarrow IR \text{ tel que } \forall x,y \text{ inclus dans } E, \text{ on a}$$

$$Msim(x, y) = \frac{\max_{p \in \text{PARING}(x,y)} \sum_{(n,n') \in p} \sigma(n,n')}{\max(|x|, |y|)}$$

avec PARING(x,y) est un ensemble de correspondances d'éléments de x vers des éléments y.

4. Méthodes sémantiques

Les méthodes sémantiques se basent principalement sur deux approches :

- La première approche repose sur les modèles de la logique : qui sont la satisfiabilité propositionnelle, la SAT modale ou les logiques de descriptions. (Giunchiglia ,2003), utilise des techniques issues de la satisfiabilité propositionnelle (SAT), qui permettent la vérification

de la validité d'un ensemble de formules propositionnelles.

- La deuxième approche regroupe les méthodes de déduction afin de déduire la similarité entre deux entités. Les techniques des logiques de description (le test de subsomption) permettent de vérifier les relations sémantiques entre les entités telles que l'équivalence (la similarité est égale à 1), la subsomption (la similarité est comprise entre 0 et 1) ou l'exclusion (la similarité est égale à 0).

Annexe 2 : les sous-ensembles flous

1. Définitions

La théorie des ensembles flous est en fait selon Zadeh (1965), un pas vers un rapprochement entre la précision des mathématiques classiques et la subtile imprécision du monde réel : un rapprochement né de l'incessante quête humaine pour une meilleure compréhension des cheminements mentaux de la connaissance (Kaufman, 1973). Elle a donc pour objet d'étude, la représentation des connaissances imprécises et le raisonnement approché. De ce fait (Gacogne, 1997), on peut la situer à côté des heuristiques de résolutions de problèmes, des systèmes experts, de l'apprentissage, de l'intelligence artificielle distribuée et même du traitement de la langue naturelle.

Les sous-ensembles flous sont utilisés pour représenter des valeurs imprécises, comme un intervalle aux limites mal définies. Ils peuvent également servir à l'expression de préférences dans les critères de sélection de requêtes en bases de données. Un exemple expressif est la proximité entre deux objets quelconques, la relation "*est proche de*" manque de précision et dépend de son utilisateur et du contexte de son application. Le concept des sous-ensembles flous constitue une solution pour le traitement des situations qui présentent des problèmes d'incertitude et d'imprécision. En effet, il permet des graduations dans l'appartenance d'un élément à une classe.

Un sous-ensemble flou est décrit sur un référentiel donné par une fonction d'appartenance à valeur sur une échelle totalement ordonnée, bornée supérieurement et inférieurement. On emploie le plus souvent en pratique l'intervalle réel $[0, 1]$, mais cette échelle peut aussi ne comporter qu'un nombre fini de niveaux. Des degrés d'appartenance intermédiaires sont ainsi autorisés entre l'appartenance totale (1) et la non-appartenance complète (0). L'idée de (sous)-ensemble flou a été introduite et élaborée par Zadeh à partir de 1965.

Dans un ensemble de référence X , depuis [Zadeh, 1965] et [Kaufmann, 1972], un sous ensemble flou de ce référentiel X est caractérisé par une fonction d'appartenance μ de X dans l'intervalle des nombres réels $[0, 1]$.

Dans une approche floue un élément peut appartenir plus au moins fortement à une classe.

2. Formulation

Un sous ensemble flou A d'un référentiel X est caractérisé par une fonction d'appartenance $\mu_A : \forall x \in X \quad \mu_A \in [0,1]$

L'ensemble A est défini par : $A = \{(x, \mu_A(x)) \mid x \in X\}$

Si $\mu_A(x) = 0.1$ signifie que x appartient à l'ensemble flou A avec un degré d'appartenance de 10%, il s'agit de la valeur de vérité de l'appartenance de x à A .
Pour un sous-ensemble flou A d'un référentiel E on donne les définitions suivantes :

3. Concepts fondamentaux

La fonction d'appartenance décrivant un sous-ensemble flou est caractérisée par les quatre propriétés suivantes :

- **Le type** : Définit la forme du nombre flou qui peut être triangulaire, trapézoïdale, gaussienne ou sigmoïdale.
- **La hauteur** : La hauteur du sous-ensemble flou A de X , notée $h(A)$, est le plus fort degré avec lequel un élément de X appartient à A . C'est la borne supérieure $h(A) = \sup_{(x \in X)} \mu_A(x)$ de la fonction d'appartenance, aussi n'est-elle pas nécessairement atteinte. Par exemple, si G comme «grand» est l'ensemble flou défini par $th(x)$ sur \mathbb{R}^+ , $th(x)$ croît de la valeur 0 en 0 vers la limite 1 quand x tend vers l'infini, mais cette limite 1 n'est jamais atteinte. Le noyau de G est l'ensemble vide, son support est \mathbb{R}^{+*} et sa hauteur 1. Un ensemble flou est dit normalisé s'il est de hauteur 1.

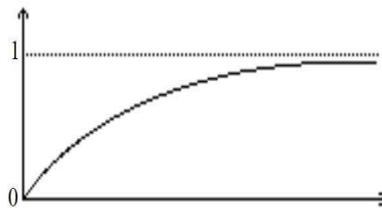


Figure A2.1 : Ensemble flou avec une hauteur qui tend vers 1

- **Le noyau** : $N(A) = \{x \mid \mu_A(x) = 1\}$, c'est l'ensemble des éléments qui sont "vraiment" dans A , autrement dit qui appartiennent totalement à A . Pour les fonctions de type triangulaire, le noyau est un singleton qui s'appelle aussi valeur modale.
- **Le support** : $S(A) = \{x \mid \mu_A(x) \neq 0\}$: Cet ensemble décrit les éléments qui sont partiellement dans A .

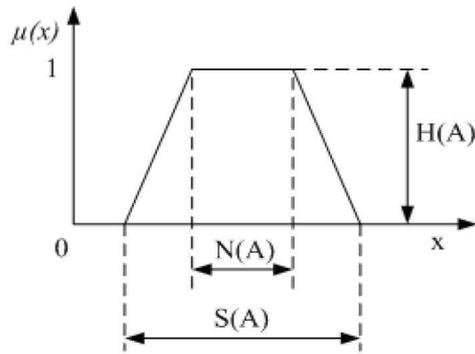


Figure A2.2 : Illustration de la hauteur, du noyau et de support d'un ensemble flou

- **Cardinalité :** La cardinalité d'un sous-ensemble flou A de X, noté $|A|$, est le nombre d'éléments appartenant à A pondéré par leur degré d'appartenance. Formellement, pour A fini :

$$\begin{cases} \mu_A(x) = 0 \\ 0 < \mu_A(x) < 1 \\ \mu_A(x) = 1 \end{cases}$$

Si A est sous-ensemble ordinaire de X, sa cardinalité est le nombre d'éléments qui le composent, selon la définition classique.

4. Les nombres flous trapézoïdaux

Un nombre flou trapézoïdal est noté généralement par un quadruplet de paramètres (a, b, α, β).

$$\mu_A(x) = \begin{cases} 0 & \text{si } x < a - \alpha \text{ ou } b + \beta < x \\ 1 & \text{si } a < x < b \\ 1 + (x - a) / \alpha & \text{si } a - \alpha < x < a \\ 1 - (b - x) / \beta & \text{si } b < x < b + \beta \end{cases}$$

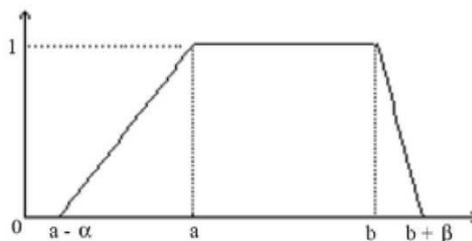


Figure A2.3 : Nombre flou trapézoïdal

5. Les nombres flous triangulaires

Un nombre flou triangulaire est cas particulier d'un nombre trapézoïdal, il est noté généralement par les trois paramètres (a, α, β) .

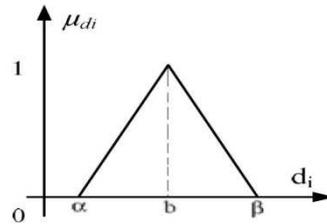


Figure A2.4 : Nombre flou triangulaire

Ils sont définis par :

$$\mu_A(x) = \begin{cases} 0 & \text{si } x < a - \alpha \\ 1 & \text{si } x = a \end{cases}$$

6. Les nombres flous avec des côtés paraboliques

Ils sont définis ainsi :

$$\mu_A(x) = \begin{cases} 0 & \text{si } x < a - \alpha \text{ ou } b + \beta < x \\ 1 & \text{si } a < x < b \\ \frac{2(x - a + \alpha)^2}{\alpha^2} & \text{si } a - \alpha < x < a - \alpha / 2 \\ 1 + \frac{2(x - a)^2}{\alpha^2} & \text{si } a - \alpha / 2 < x < a \\ -\frac{2(b - x)^2}{\beta^2} & \text{si } b < x < b + \beta / 2 \\ \frac{2(x - b + \beta)^2}{\beta^2} & \text{si } a + \beta / 2 < x < b \end{cases}$$

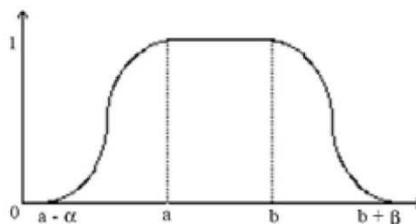


Figure A2.5 : Nombre flou avec des côtés paraboliques

7. Distance entre sous ensembles flous

Plusieurs notions de distances entre sous-ensembles flous ont été proposées dans la littérature. Nous citons ci-après les distances de Hamming, euclidienne, de Kosko et de Bhattacharyya.

Dans les formules suivantes, A et B sont deux sous-ensembles flous d'un référentiel X.

a. Distance de Hamming

La distance de Hamming entre deux sous-ensembles flous A et B est une mesure qui indique le degré global avec lequel les éléments de μ appartiennent à A et/ou B. Elle peut être utile pour définir des relations de type « à peu près égal » ou « très supérieur à ». Cette mesure permet d'évaluer un index de similarité plutôt qu'une mesure de proximité de deux ensembles.

La distance de Hamming est : $d_{hm}(A, B) = \sum_{x \in X} |\mu_A(x) - \mu_B(x)|$, c'est aussi l'intégrale

$$d_{hm}(A, B) = \int_a^b |\mu_A(x) - \mu_B(x)| dx \text{ sur l'ensemble des réels } \square .$$

En outre, si le référentiel est fini, il peut être intéressant de définir une distance relative par le rapport $\delta(A, B) = \frac{d(A, B)}{\text{card}(X)}$ qui sera maximal égal à 1 entre les parties vide et pleine.

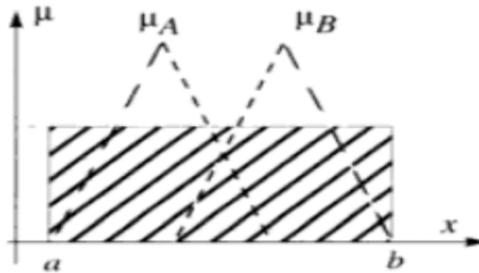


Figure A2.6 : Distance de Hamming relative

b. Distance euclidienne

Cette distance mesure la ressemblance entre deux sous-ensembles flous, elle est

définie ainsi : $d_e(A, B) = \sqrt{\sum_{x \in X} (\mu_A(x) - \mu_B(x))^2}$

La distance euclidienne pondérée est comparable à la distance euclidienne mais permet l'introduction de coefficient W_i pour donner plus d'importance à la différence pour un x_i donné. Les coefficients W_i permettent également de normaliser le résultat de cette distance entre 0 et 1. Leur choix dépend de l'application et ils sont fixés à la suite d'un ensemble d'essais. Cette distance est définie ainsi :

$$d_{ep}(A, B) = \sqrt{\sum_{x \in X} W_i^2 (\mu_A(x) - \mu_B(x))^2}$$

c. Distance de Kosko

Cette distance est une généralisation de la distance Euclidienne. Elle est définie ainsi :

$$d_k(A, B) = \sqrt[p]{\sum_{x \in X} |\mu_A(x) - \mu_B(x)|^p}$$

Le paramètre p est un réel appartenant à l'intervalle $[0,4]$.

d. Distance Bhattacharyya

Elle permet d'évaluer la similarité sémantique entre deux sous ensembles flous et par conséquent. Cette distance peut être utilisée lorsque les limites de la distance euclidienne auront été atteintes. Par exemple pour des vocabulaires où les fonctions de références sont mal séparées. Cependant, elle nécessite un temps de calcul plus long que la distance euclidienne.

La distance de Bhattacharyya sera utilisée pour des applications où il est nécessaire d'avoir un bon compromis entre la vitesse et la précision.

Soient :

$$p_A(x \in X) = \frac{\mu_A(x)}{\sum \mu_A(x)} \text{ et } p_B(x \in X) = \frac{\mu_B(x)}{\sum \mu_B(x)}$$

Le coefficient R de Bhattacharyya est défini ainsi : $R(p_A, p_B) = \sum_{x \in X} \sqrt{p_A(x) \cdot p_B(x)}$. Ce

coefficient varie entre 0 et 1. L'égalité avec l'unité est atteinte lorsque A et B sont identiques. La distance de Bhattacharyya est comprise entre 0 et 1 et est définie par :

$$d_{bh}(A, B) = \sqrt{1 - R(p_A, p_B)}$$

Annexe 3

Convention de cotutelle

**CONVENTION
DE CO-TUTELLE DE THESE
FRANCO-TUNISIENNE**

Entre

**L'ECOLE NATIONALE SUPERIEURE DES
TELECOMMUNICATIONS DE BRETAGNE
(TELECOM Bretagne)**

Et

**L'ECOLE NATIONALE DES SCIENCES DE
L'INFORMATIQUE
(ENSI TUNISIE)**

ENTRE LES SOUSSIGNES :

L'Ecole Nationale Supérieure des Télécommunications de Bretagne (France)

Sise Technopôle de Brest-iroise, CS 83818, 29238 BREST Cedex

représentée aux fins des présentes par Monsieur André CHOMETTE, Directeur de l'Ecole Supérieure des Télécommunications de Bretagne et ci-après dénommée **TELECOM Bretagne**

ET

L'Ecole Nationale des Sciences de l'Informatique (Tunisie)

Relevant de l'Université de la Manouba, sise au Campus Universitaire de Manouba, 2010 Manouba, représentée aux fins des présentes par Mme Leila SAIDANE, Directeur de l'Ecole Nationale des Sciences de l'Informatique.

IL EST D'ABORD EXPOSE CE QUI SUIIT :

VU la loi n° 84-52 du 26 janvier 1984 sur l'enseignement supérieur français,

VU l'arrêté du 21 août 2000 fixant la liste des établissements d'enseignement supérieur autorisés à délivrer le doctorat conjointement avec une université ou un institut national polytechnique

VU l'arrêté du 6 janvier 2005 relatif à la co-tutelle internationale de thèse

VU l'arrêté du 7 août 2006 relatif à la formation doctorale

VU le décret n° 97-1801 du 3 septembre 1997, modifiant et complétant le décret n° 93-1823 du 6 septembre 1993, fixant les conditions d'obtention des diplômes nationaux sanctionnant les études doctorales.

Les deux parties, animées par la volonté de favoriser les échanges des doctorants entre elles et de renforcer, ainsi, la coopération scientifique et universitaire entre la France et la Tunisie, décident d'un commun accord, dans le cadre de la législation en vigueur dans leurs pays respectifs, d'utiliser la procédure de co-tutelle concernant :

M. Wassim MESSOUDI de nationalité Tunisienne

Adresse en Tunisie : 117 Cité Elons, Jendouba Nord 8189, Tunisie

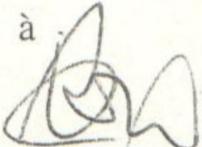
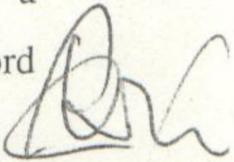
Adresse en France : TELECOM Bretagne, Département

Technopôle Brest Iroise CS 83818 – Brest 29200

CECI ETANT EXPOSE, IL A ETE CONVENU ET ARRETE CE QUI SUIV

Article 1

Le Doctorant doit être inscrit dans les deux établissements. Il doit s'acquitter des frais d'inscription selon les modalités suivantes :

1. *Première* inscription de thèse : le Doctorant doit s'acquitter des droits d'inscription à l'Ecole Nationale des Sciences de l'Informatique (ENSI).
2. *Deuxième* inscription de thèse : le Doctorant doit s'acquitter des droits d'inscription à TELECOM Bretagne et à l'Ecole Nationale des Sciences de l'Informatique (ENSI): 
3. *Troisième* inscription de thèse : le Doctorant doit s'acquitter des droits d'inscription à TELECOM Bretagne et à l'Ecole Nationale des Sciences de l'Informatique (ENSI), sauf accord signé entre les deux pays. 
4. Le Doctorant sera exonéré de ces mêmes droits à l'établissement partenaire sur présentation d'une attestation d'inscription ou d'un justificatif délivré par l'établissement où les droits d'inscription ont été acquittés.

Dans ce contexte, les deux parties prennent acte et enregistrent les données suivantes :

1. Date de l'inscription en thèse sous le régime de co-tutelle : 24 Janvier 2009
2. Titre de la thèse : Vers une approche de recherche stratégique d'images satellitales par le contenu sémantique basée sur la fusion ontologique
3. Durée prévisionnelle des travaux de Recherche : **3ans** (A compter du 24 Janvier 2008)
(date prévisionnelle de soutenance : 24 Janvier 2011)
 - périodes prévues à TELECOM Bretagne : Du Mai à Aout 2009 et 2010
 - périodes prévues à ENSI : de Janvier à Avril et Septembre à Décembre 2009 et 2010

Article 2

Lors de son séjour à TELECOM Bretagne, la couverture sociale du doctorant est assurée par lui-même conformément à la législation en vigueur, il se charge aussi de son hébergement en résidence.

Article 3

Dans chacun des établissements concernés, le doctorant effectuera ses travaux de recherche sous la direction et la responsabilité des Directeurs de thèse suivants :

Pour TELECOM Bretagne : Prof. Basel SOLAIMAN
Pour ENSI Tunisie : Prof. Henda BEN GHEZALA

Ceux-ci s'engagent à exercer pleinement et conjointement, auprès du doctorant, les compétences qui leur sont attribuées par la réglementation en vigueur et les traditions universitaires dans leurs pays respectifs.

Article 4

La composition du jury de soutenance obéit à la réglementation en vigueur dans les deux pays.

Le jury sera composé à parité de chaque côté de représentants scientifiques des deux pays. Il comprend obligatoirement les deux Directeurs de thèse et un membre extérieur aux deux établissements.

Le jury de soutenance devra être désigné trois mois avant la date effective de soutenance.

Article 5

La thèse, préparée en co-tutelle, sera rédigée en *français*, et complétée par un résumé écrit en *anglais*.

Article 6

La thèse donnera lieu à une soutenance unique en *France*, reconnue par les deux établissements.

La délivrance des deux diplômes :

Sur avis favorable du jury de soutenance,

TELECOM Bretagne en habilitation conjointe avec l'Université Rennes I, s'engage à délivrer à M Wassim MESSOUDI le grade de docteur et à lui délivrer le diplôme correspondant.

Et

L'Ecole nationale des sciences de l'informatique ENSI Relevant de l'Université de la Manouba, s'engage à délivrer à M Wassim MESSOUDI le grade de docteur et à lui délivrer le diplôme correspondant.

Le libellé de chaque diplôme fera mention de la collaboration de l'établissement partenaire ainsi que de la cotutelle.

Article 7

Les modalités de dépôt, signalement et reproduction de la thèse ainsi que l'autorisation de soutenir obéissent à la réglementation en vigueur dans les deux pays.

La date et le lieu de soutenance sont fixés d'un commun accord et notifiés par écrit par les co-directeurs de thèse aux deux chefs d'établissement.

Article 8

La protection du sujet de thèse ainsi que la publication, l'exploitation et la protection des résultats de recherche issus des travaux du doctorant sont assujetties à la réglementation en vigueur dans les deux établissements conformément aux procédures spécifiques à chaque pays impliqué dans la cotutelle.

Les dispositions relatives à la protection des droits de propriété intellectuelle pourront faire l'objet d'une convention spécifique, le cas échéant.

Article 9

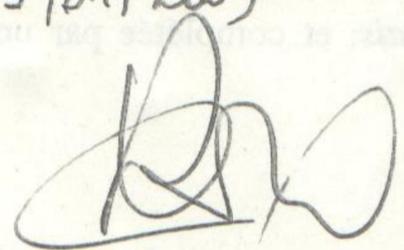
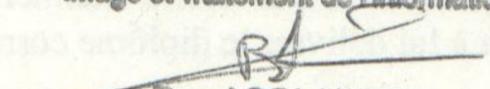
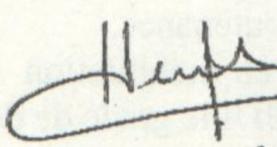
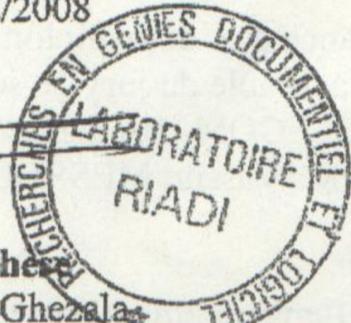
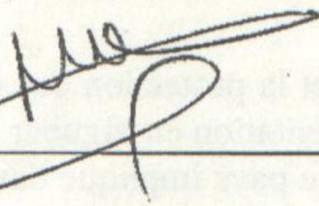
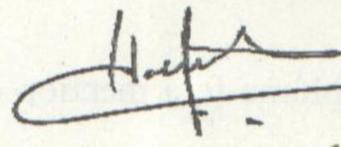
Soucieux de l'intérêt des doctorants et du développement de la coopération entre eux et leurs pays respectifs, les Etablissements d'accueil s'engagent à respecter les dispositions ci-dessus et à faire tout ce qui est nécessaire pour l'application de la présente convention dans les meilleures conditions.

En cas de litige, les parties à la présente convention s'engagent à rechercher toute solution amiable avant d'en décider la résolution.

Article 10

Au cas où le régime de co-tutelle viendrait à être dénoncé par une des parties concernées, celle-ci devra le notifier à l'autre partie par écrit dans le délai d'un mois en indiquant les raisons de sa décision.

Fait en sept exemplaires originaux.

<p>Pour TELECOM Bretagne</p>	<p>Pour l'Ecole Nationale des Sciences de l'Informatique (ENSI)</p>
<p>A Brest, le 19/01/2009</p>  <p>Le Directeur M. André CHOMETTE</p>	<p>A Tunis, le 24/12/2008</p>   <p>Le Directeur Mme Leila Saidane</p>
<p>A Brest, le</p> <p>Le Chef du Département Image et Traitement de l'Information</p>  <p>Basel SOLAIMAN</p> <p>Le Directeur de thèse M. Basel SOLAIMAN</p>	<p>A Tunis, le 24/12/2008</p>   <p>Le Directeur de thèse Mme. Henda Ben Ghezala</p>
<p>A Tunis, le 24/12/2008</p> <p>Le Doctorant Wassim MESSOUDI</p> 	<p>A Tunis, le 24/12/2008</p>   <p>Le Président de l'Université de la Manouba Mme. Henda Ben Ghezala</p>

Publications

Chapitre de livre (1) :

"Modélisation de l'information géographique par ontologie spatio-temporelle, floue et lourde pour l'interprétation des images de télédétection", Wassim Messaoudi, Imed Riadh Farah and Basel Solaiman, à paraître dans l'ouvrage «*Information géographique en Afrique, Etat des lieux et nouvelles approches*»

Journal (1) :

"Satellite Image Retrieval Based on Ontology Merging", Imed Riadh Farah, Wassim Messaoudi, Karim saheb ettabâa, Basel Solaiman, ICGST-GVIP Journal, ISSN: 1687-398X, Volume 8, Issue 2, July 2008, pp 45-53.

Soumis (2) :

"Fuzzy, spatio-spectro-temporal, heavy and evolutionary ontology for remote sensing image annotation and reasoning; Application to natural risks management", Wassim Messaoudi, Imed Riadh Farah and Basel Solaiman, (*Soumis au journal : Applied Ontology*)

"Vers une ontologie riche dédiée à l'imagerie satellitaire par réutilisation de ressources existantes", Hafedh Nefzi, Wassim Messaoudi, Mohamed Farah, Imed Riadh Farah, Basel Solaiman (*Soumis à Taima'13*)

Conférences internationales (5) :

"Vers une ontologie spatio-temporelle, floue et lourde de l'imagerie satellitale", Wassim Messaoudi, Imed Riadh Farah and Basel Solaiman, 4èmes Journées Francophones sur les Ontologies, 22 – 23 Juin 2011, Montréal, Canada

"A methodology for modelling and retrieving satellite images basing on spatial knowledge: application to natural risks", Wassim Messaoudi, Imed Riadh Farah, Henda Ben Ghezala and Basel Solaiman, IEEE International Conference on Graphic and Image Processing, Manila, Philippines. December 4-5, 2010.

"Proposition d'une méthodologie basée sur la représentation de connaissances de l'image satellitale pour la prévention contre les risques naturels : Application à l'érosion", Wassim Messaoudi, Imed Riadh Farah, Henda Ben Ghezala and Basel Solaiman, XII^{ème} journées scientifiques du réseau Télédétection de l'A.U.F, Monastir, Tunisie, 23-25 Novembre 2010.

"Spatio-temporal multi-modality ontology for indexing and retrieving satellite images", Wassim Messaoudi, Imed Riadh Farah, Karim saheb ettabâa, Henda Ben Ghezala and Basel Solaiman, 6^{ème} Colloque sur l'Optimisation et les Systèmes d'Information (COSI'2009), 25-27 Mai 2009, Annaba Algérie.

"Semantic strategic satellite image retrieval", W. Messaoudi, I.R. Farah, K. Saheb-Ettabâa, B. Solaiman, 3rd IEEE International Conference on Information and Communication Technologies : from theory to application (ICTTA'08), 7-11 April 2008, Damascus Syria.

Bibliographie

Baader F., Hollunder B., "A Terminological Knowledge Representation System with Complete Inference Algorithms", Proceedings of the Workshop on Processing Declarative Knowledge, 1991.

Bach P., "Construction d'un Web sémantique multi-points de vue", thèse de doctorat, Ecole des Mines de Paris à Sophia Antipolis, soutenue le 23 octobre 2006.

Bachimont B., "Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances", Ingénierie des connaissances : évolutions récentes et nouveaux défis, p. 305-323, Eyrolles, 2000.

Ball G. H., Hall, D. J., "A novel method of data analysis and pattern classification", Technical report, Stanford Research Institute, Menlo Park, 1965.

Bergmann R. et Stahl A., "Similarity measures for object-oriented case representations", Proceedings of the European Workshop on Case-Based Reasoning, EWC-BR'98, 1998.

Bonn F., Rochon G., "Précis de télédétection, Principes et Méthodes", Tome 1, Presses de l'Université de Québec/APELE, 1992.

Borst W. N. "Construction of Engineering Ontologies", Center for Telematica and Information Technology, University of Twente, Enschede, NL, 1997.

Bradley N., "The {XML} Companion", Addison-Wesley Professional Publisher, 2001.

Breen C., Khan L., Ponnusamy A., "Image Classification Using Neural Networks and Ontologies", 13th International Workshop on Database and Expert Systems Applications, p. 98-102, 2002.

Carbonell J. G., "Derivational analogy and its role in problem solving". AAAI-83, p. 64-69, 1983.

Charlet J., Szulman S., Aussenac-Gilles N., Nazarenko A., Hernandez N., Nadah N., Sardet E., Del Ahouse J., Teguiak V., "DAFOE: an ontology building platform from text or thesauri", 8^{ème} Rencontres Terminologie & Intelligence Artificielle, Toulouse, 18-20 Novembre 2009.

Chouraqui E., "Le raisonnement analogique : sa problématique, ses applications", Actes des Journées Nationales sur l'Intelligence Artificielle, Aix-les-Bains, CEPADUES-Editions, Toulouse, p. 107-117, 1986.

Clarke, B., "A calculus of individuals based on connection", Notre Dame Journal of Formal Logic, vol. 22, n° 3, p. 204-218, 1981.

Cullot N., Parent C., Spaccapietra S, Vangenot C, "Des ontologies pour données géographiques", Revue Internationale de Géomatique, numéro spécial : Les SIG sur le Web, vol. 13, n° 3, p. 285-306, Hermès Science Edition, 2003.

D'Aquin M., Lieber J. et Napoli A., "Decentralized Case-Based Reasoning for the Semantic Web", Proceeding of International Semantic Web Conference, ISWC 2005, tome 3729 de Lecture Notes in Computer Science, p. 142-155. Springer, 2005.

Decker S., Erdmann M., Fensel D., Horrocks I., Klein M., Harmelen F. V., "Oil in a nutshell", In Proceedings of the 12th European Knowledge Acquisition Workshop (EKAW'00), Lecture Notes in Artificial Intelligence, n° 1937, p. 1-16, 2000.

Durand N., Derivaux S., Forestier G., Wemmert C., Gançarski P., Boussaid O., Puissant A., "Ontology-based Object Recognition for Remote Sensing Image Interpretation", 19th IEEE International Conference on Tools with Artificial Intelligence, p. 472-479, 2007.

Euzenat J. et Shvaiko P., "Ontology Matching", Springer-Verlag, Heidelberg(DE), 333 pages, 2007

Freeman J., "The modelling of spatial relations", Computer Graphics and Image Processing, vol. 4, p. 156-171, 1975.

Gacogne L. "Research of Pareto set by genetic algorithm, application to multicriteria optimization of fuzzy controller", EUFIT 1997.

Guarino N., "Some organizing principles for a unified top-level ontology". AAAI Spring Symposium on Ontological Engineering, p. 57-63, 1997.

Guarino N., Welty, C. "Ontological analysis of taxonomic relationships", Proceedings of 19th International Conference on Conceptual Modeling, Salt Lake City USA, 2000.

Girard M.C et Girard C., "Traitement de données de télédétection", Dunod, Paris, 1999.

Giunchiglia F., "S-match : an algorithm and an implementation of semantic matching", In proceedings of the 1st european semantic web symposium(ESWS), vol. 3053 of LNCS. Springer-Verlag p. 61-75, 2003.

Goetz A., "Imaging spectrometry: sensors and data analysis", Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS, Houston TX, USA, 1992.

Gruber T.R. "A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition", vol. 5, n° 2, p. 199-220, 1993.

Grüniger M., Fox. "M.S. Methodology for the Design and Evaluation of Ontologies", Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal, 1995.

Heijst G. V., Schreiber G., Wielinga B., "Using explicit ontologies for KBS development", International Journal of Human-Computer Studies, vol. 42, p. 183-292, 1997.

Horrocks I, "DAML+OIL: a Description Logic for the Semantic Web", IEEE Data Engineering Bulletin, vol. 25, p. 4-9, 2002.

Hudelot C., Maillot N., and Thonnat M., "Symbol Grounding for Semantic Image Interpretation: From Image Data to Semantics". In Proceedings of the Tenth IEEE international Conference on Computer Vision Workshops, Beijing, China, October 17-20, 2005.

Hudelot C., Atif J., Bloch I. "Ontologie de relations spatiales floues pour l'interprétation d'images", In Rencontres francophones sur la Logique Floue et ses Applications, LFA-2006, Toulouse (France), 2006.

Ji. M., "Using fuzzy sets to improve cluster labelling in unsupervised classification", International Journal of Remote Sensing, vol. 24, n°4, p. 657-671, 2001.

Johaniyak, Z.C., kovacs S., "Distance based similarity measures on fuzzy sets", Proceeding of the Third Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence, SAMI-2005, Herl'any, Slovakia, 2005.

Kaufmann A. "Introduction à la théorie des sous ensembles flous", tome 1, Masson, 1973.

Kiryakov A., Popov B., Ognyanoff D., Manov D., Goranov K. M., "Semantic annotation, indexing, and retrieval", Journal of Web Semantics, vol. 2, p. 49-79, 2004.

Lassila O., Swick R. R., "Resource description framework (rdf) model and syntax", W3C Recommendation, World Wide Web Consortium, Cambridge (MA), February 1999.

Lassila O., McGuinness D., "The role of frame-based representation on the semantic Web", Computer and Information Science, vol. 6, n° 5, Linköping University, 2001

Lukasiewicz T., "Fuzzy Description Logic Programs under the Answer Set Semantics for the Semantic", Fundamenta Informaticae, vol. 8, n° 3, p. 289-310, 2008.

Mechouche A., Morandi X., Golbreich C., Gibaud B., "A Hybrid System Using Symbolic and Numeric Knowledge for the Semantic Annotation of Sulco-Gyral Anatomy in Brain MRI Images", IEEE Transactions on Medical Imaging, vol. 28, n° 8, p. 1165-1178, 2009.

Messaoudi W., Farah I.R., Ben Ghezala H., Solaiman B., "A methodology for modelling and retrieving satellite images basing on spatial knowledge: application to natural risks", Proceedings of the SPIE International Conference on Graphic and Image Processing (ICGIP 2010), vol. 8285, 2010.

Mezaris V., Kompatsiaris I., Strintzis M., "An ontology approach to object-based image retrieval", Proceedings of the IEEE International Conference on Image Processing, ICIP03, vol. 2, p. 511-514, 2003.

Miller G.A., "Wordnet : alexical Database for English". Communication of the ACM, vol. 38, n°11, p. 39-41, 1995.

Minsky M., "A framework for representing knowledge", Psychology of Computer Vision, P.H. Winston (Ed), p. 211-277, 1975.

Mizoguchi R. et Bourdeau J., "Theory-Aware Authoring Environment: Ontological Engineering Approach", Proceedings of the ICCE Workshop on Concepts and Ontologies in Web-based Educational Systems, Technische Universiteit Eindhoven, 2002.

Mustière S., Abadie N., Aussenac-Gilles N., Bessagnet M.-N., Kamel M., Kergosian E., Reynaud C., Safar B., "GéOnto : Enrichissement d'une taxonomie de concepts topographiques", Colloque International de Géomatique et d'Analyse Spatiale SAGEO'09, Paris (France), 25-27 novembre 2009.

Natalya F., Noy D., Mcguinness. L. "Développement d'une ontologie 101 : Guide pour la création de votre première ontologie", Université de Stanford, Stanford, CA, 94305, USA, 2002.

Neeches R., Fikes R. E., Finin T., Gruber T. R., Senator T., Swartout W. R. "Enabling technology for knowledge sharing", AI Magazine, vol. 12, n° 3, 1991.

Noy N.F., Musen M. A., "The prompt suite: interactive tools for ontology merging and mapping", International Journal of Human-Computer Studies, vol. 59, n° 6, p. 983-1024, 2003.

Noy N. F., Musen M. A., "Smart: automated support for ontology merging and alignment", Proceedings of the Twelfth Banff Workshop on Knowledge Acquisition, Modeling, and Management, Banff, Alberta 1999.

Paslaru-Bontas E., Schlangen D., Schrader T., "Creating Ontologies for Content Representation - the OntoSeed Suite", Proceedings of the 4th International Conference on Ontologies, Databases, and Applications of Semantics", p. 1296-1313, 2005.

Pearlman J., Caramen S., Segal C., Jarecke P., Clancy P. et Browne W., "Overview of Hyperion image spectrometer for the NASA EO-1 mission", International Symposium On Geoscience and Remote Sensing, tome 7, p. 3036-3038, 2001.

Pinto, H.S., A. Gómez-Pérez, and J.P. Martins. "Some Issues on Ontology Integration", Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends, Stockholm (Sweden), 1999.

Puissant A., Sheeren D., Weber C., Gańczarski P., Wemmert C., "Amélioration des connaissances sur l'environnement urbain : intérêt de l'intégration de règles et de l'utilisation de classifications multi-formalisme", Actes du colloque Interactions Nature-Société : analyses et modèles, La Baule (France), 3-6 mai 2006.

Quillian M.R., "Semantic memory", Semantic Information Processing, p. 216-270, MIT press, Cambridge, 1968.

Rahm E., Bernstein P., "A survey of approaches to automatic schema matching", VLDB Journal, vol. 10, n° 4, p.334-350, 2001.

Roussey C., Laurini R., Beaulieu C., Tardy Y., Zimmerman M., "Le projet Towntology: un retour d'expérience pour la construction d'une ontologie urbaine", Revue Internationale de Géomatique, vol. 14, n° 2, pp. 217-237, 2004.

Ruan, N. Huang, N. Hong, W. "Semantic-Based Image Retrieval in Remote Sensing Archive: An Ontology Approach", Geoscience and Remote Sensing Symposium, IGARSS, 2006.

Shvaiko P. et Euzenat J., "A survey of schema-based matching approaches", Journal on data semantics, vol. 4, p.146-171, 2005.

Smeulders A. W., Worring M., Santini S., Gupta A., Jain R. "Content-based image retrieval at the end of the early years", IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 22, n° 12, p. 1349-1380, 2000.

Smyth B., Keane M. T., Cunningham P., "Hierarchical case-based reasoning integrating case-based and decompositional problem-solving techniques for plant-control software design", IEEE Transactions on Knowledge and Data Engineering, vol. 13, n° 5, p. 793-812, 2001.

Sowa J.F., "Conceptual Structures: Information Processing in Mind and Machine", Addison-Wesley Publishing Company, USA, 1984.

Sowa J.F., "Top-level ontological categories", International Journal of Human and Computer Studies, vol. 43, p. 669-685, 1995.

Staab S., Maedche A., "Axioms are objects too: Ontology engineering beyond the modeling of concepts and relations", Research report 399, Institute AIFB, Karlsruhe, 2000.

Studer R., Benjamins V. R. et Fensel D. "Knowledge engineering: Principles and Methods" Data Knowledge Engineering, vol. 25, n° 1-2, 1998.

Swartout B., Patil R., Knight K. et Russ T. "Towards Distributed Use of Large Scale Ontologies", Spring Symposium Series on Ontological Engineering, Stanford University, CA, 1997.

Uschold M., Grüninger M., "Ontologies: Principles, Methods and Applications", Journal of Knowledge Engineering Review, vol. 11, p. 93-136, 1996.

Uschold M., King M., "Towards a methodology for building ontologies", Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95, 1995.

Valtchev P., "Construction automatique de taxonomies pour l'aide à la représentation de connaissance par objets", thèse de doctorat, Université de Grenoble 1, France, soutenue le 15 Décembre 1999.

Wang H., Liu S., Chia L. T., "Does ontology help in image retrieval ? - a comparison between keyword, text ontology and multi-modality ontology approaches", Proceedings of the 14th annual ACM international conference on Multimedia, p. 109-112, New York, NY, USA, 2006.

Wu Z. et Palmer M., "Verbs Semantics and Lexical Selection", Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, p. 133 - 138 , New Mexico, USA, 1994.

Yen J., Langari R., "Fuzzy logic: Intelligence, control, and information", Prentice Hall, 1999.

Zadeh L.A., "Fuzzy Sets", Information Control, Vol. 8, p. 338-353, 1965.