



HAL
open science

Approches multi-hiérarchiques pour l'analyse d'images de télédétection

Camille Kurtz, Pierre Gañarski, Anne Puissant, Nicolas Passat

► **To cite this version:**

Camille Kurtz, Pierre Gañarski, Anne Puissant, Nicolas Passat. Approches multi-hiérarchiques pour l'analyse d'images de télédétection. *Revue Française de Photogrammétrie et de Télédétection*, 2014, 205, pp.19-35. hal-01694404

HAL Id: hal-01694404

<https://hal.univ-reims.fr/hal-01694404>

Submitted on 3 Mar 2018

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.

APPROCHES MULTI-HIÉRARCHIQUES POUR L'ANALYSE D'IMAGES DE TÉLÉDÉTECTION

Camille Kurtz¹, Pierre Gançarski², Anne Puissant³, Nicolas Passat⁴

1: Université Paris Descartes, LIPADE (camille.kurtz@parisdescartes.fr)

2: Université de Strasbourg, CNRS, ICube (gancarski@unistra.fr)

3: Université de Strasbourg, CNRS, LIVE (anne.puissant@live-cnrs.unistra.fr)

4: Université de Reims Champagne-Ardenne, CReSTIC (nicolas.passat@univ-reims.fr)

Résumé

Au cours de la dernière décennie, la télédétection par imagerie satellitaire a connu plusieurs révolutions, liées d'une part à l'augmentation spectaculaire du nombre de dispositifs d'acquisition, et d'autre part aux progrès des capteurs tant en termes de résolutions spatiale et temporelle que spectrale. Dans ce contexte, de nouvelles problématiques sont apparues, notamment en traitement et analyse d'images. L'un des axes de recherche les plus prometteurs pour traiter ces problématiques s'appuie sur la notion de hiérarchie, qui peut se décliner sous plusieurs formes et ainsi permettre de traiter plusieurs types de tâches, du traitement de bas-niveau jusqu'à l'analyse de haut-niveau. Dans cet article, nous décrivons certaines tendances récentes liées à ces approches hiérarchiques.

Mots clés : segmentation, classification, hiérarchies, morphologie mathématique

Abstract

During the last decade, remote sensing from space has experienced several revolutions, due on the one hand to the dramatic increase of the number of image acquisition devices, and on the other hand to the progress of sensors in terms of spatial, temporal and spectral resolutions. In this context, new issues have been raised, in terms of image processing and analysis. One of the most promising research axes to tackle these issues relies on the notion of hierarchy, that can be declined under several forms, to deal with various tasks, from low-level processing to high-level analysis. In this article, we describe some recent trends related to these hierarchical approaches.

Keywords : segmentation, classification, hierarchy, mathematical morphology

1. Contexte

1.1. De nouvelles images...

Ces dernières années, les images de télédétection ont fait l'objet de multiples révolutions technologiques.

La première d'entre elles est – à double titre – une révolution de masse. D'une part, car elle a porté sur la surmultiplication des données images acquises, notamment par l'augmentation du nombre de capteurs à disposition. La mise en place de constellations de satellites – par exemple, le programme « Pléiades » du CNES (de Lussy et al., 2005) ou encore le programme européen « Copernicus »¹ – illustre cette augmentation. D'autre part, car cette inflation du nombre d'images satellitaires s'intègre dans un mouvement plus général de globalisation et d'accessibilité des informations (internet, télécommunications, géolocalisation, etc.), mouvement qui impacte directement le quotidien des populations, et dont les implications économiques et sociétales sont multiples.

La seconde révolution est dimensionnelle. L'accroissement du nombre de capteurs s'accompagne de l'apparition d'informations supplémentaires, au delà de la dimension spectrale. L'une de ces nouvelles dimensions

est temporelle. Il est ainsi désormais aisé d'obtenir des séquences d'images, et ce à une fréquence élevée (quelques jours), là où les données étaient auparavant considérées comme statiques (annuelle). Les nouvelles dimensions sont également spatiales, puisque les progrès accomplis, par exemple dans les modalités RADAR (Krieger et al., 2008) et LiDAR (Liu, 2008), permettent désormais de compléter des informations jusque là radiométriques par des données géométriques et morphologiques. à cela s'ajoute une capacité croissante à intégrer des informations sémantiques (Sheth et al., 2008), notamment par le développement de nouvelles technologies, par exemple liées à l'internet des objets et à la géolocalisation.

La troisième et dernière révolution – mais néanmoins pas la moindre – concerne plus classiquement les progrès réalisés en imagerie optique. Outre une amélioration progressive de la résolution spectrale des satellites (e.g., imagerie hyper-spectrale (Goetz, 2009)), le principal progrès concerne l'amélioration de la résolution spatiale des images (Benediktsson et al., 2012). En quelques années, les images – dans le domaine civil – sont ainsi passées de « moyennes résolutions » (MR, 30–10 mètres) à des « hautes résolutions » (HR, 5–1

1. URL : <http://www.copernicus.eu>

mètre(s)) pour atteindre désormais des « très hautes résolutions » (THR, inférieures à 1 mètre). L'apparition de ces nouvelles résolutions n'a néanmoins pas induit la disparition des capteurs à des résolutions moindres, et les images mises à disposition s'échelonnent dorénavant dans des échelles spatiales particulièrement vastes.

1.2. ... pour de nouveaux défis

Ces révolutions technologiques s'insèrent dans un cadre beaucoup plus général, d'explosion de la quantité des informations, de diversification de leurs sources, et d'hétérogénéisation de leurs contenus. Ces différentes inflations – qui touchent tous les domaines, au delà de la seule télédétection – induisent de nouvelles problématiques liées à la gestion et à l'extraction des connaissances, dans un contexte qui a donné lieu à l'apparition du concept de « big data » (Jacobs, 2009).

à une moindre échelle, et dans le contexte thématique de cet article, la problématique considérée porte sur l'analyse des images de télédétection. Dans ce cadre plus restreint, les défis restent toutefois nombreux. En effet, le changement de nature des images de télédétection rend désormais obsolète une conception purement psychovisuelle de leur interprétation. En d'autres termes, l'œil – et le cerveau – de l'expert humain ne fournissent plus des outils suffisants, tant qualitativement que quantitativement, pour une analyse efficace.

D'un point de vue pratique, la taille des images rend désormais une analyse par photo-interprétation impraticable, et motive le développement de méthodes automatiques, ou a minima faiblement interactives. Cette nécessité d'assister l'expert par des outils logiciels est également justifiée par l'apparition dans les images à très hautes résolutions d'une multitude de détails qui peuvent tout aussi bien contribuer à la richesse de l'information cherchée qu'à l'apparition de bruit sémantique.

D'un point de vue analytique, la disponibilité d'images multi-résolutions permet également d'observer une scène à plusieurs échelles, et donc à plusieurs niveaux d'interprétation. La mise en relation de ces différents niveaux est néanmoins complexe, et nécessite là encore une assistance pour l'expert, afin de gérer la mise en relation de différents schémas de représentation.

Un changement de paradigme d'utilisation et d'interprétation des images devient donc inévitable. L'expert doit toutefois rester au cœur du système, car il apporte une connaissance dont la complexité, la richesse et la versatilité se prêtent difficilement à une modélisation purement numérique. Cette hypothèse est notamment renforcée par la variabilité des objectifs thématiques en télédétection : développement urbain, gestion de catastrophes, évolution des milieux écologiques, etc.

Ce changement de paradigme touche tous les niveaux d'analyse. Le premier concerné, au niveau le plus bas, porte sur l'extraction des primitives de base des images, c'est-à-dire la segmentation de l'image en régions. à un niveau intermédiaire, le groupement de ces primitives en familles homogènes, *via* des mécanismes de classification est également impacté. Enfin, au niveau

le plus élevé, l'interprétation des images en termes de concepts sémantiques doit aussi être reconsidéré.

L'une des principales réponses apportées à ces problématiques au cours des dernières années réside dans le concept de hiérarchie, qui vise à organiser de manière cohérente des entités (spatiales, conceptuelles, sémantiques, ...) selon différents niveaux d'analyse. La hiérarchisation vise à permettre d'une part la prise en compte des informations de nature hétérogène fournies par les images, et d'autre part à autoriser l'expert humain à se focaliser plus aisément sur un niveau de hiérarchie donné, en fonction de ses besoins. La notion de hiérarchie permet également de proposer un cadre unifiant qui regroupe non seulement les informations extraites des données images, mais également celles fournies par d'autres sources, dont les connaissances de l'expert.

2. Principaux paradigmes hiérarchiques

Le concept de hiérarchie s'interface avec la notion d'images à plusieurs niveaux, depuis leur expression spatiale et spectrale, jusqu'à leur sémantique. Ces paradigmes hiérarchiques induisent alors des méthodologies de traitement, d'analyse et d'interprétation dédiées. Dans cette section, nous les décrivons du plus bas au plus haut niveau.

2.1. L'image en tant que hiérarchie

D'un point de vue structurel, une image peut être considérée comme une hiérarchie. Cette vision est notamment valable pour les images mono-valuées, c'est-à-dire à niveaux de gris. Plus précisément, chaque niveau de gris, du plus bas au plus élevé (ou inversement) est alors vu comme un niveau de la hiérarchie. Chaque « coupe » de l'image, obtenue par seuillage, s'organise hiérarchiquement par rapport à ses coupes voisines, de la plus étendue à la plus restreinte.

Cette conception hiérarchique des images a été largement considérée en morphologie mathématique (Najman et Talbot, 2008), domaine lui-même très tôt impliqué dans le développement de méthodologies de traitement et d'analyse d'images de télédétection (Soille et Pesaresi, 2002). Dans ce contexte, deux approches ont notamment été développées.

La première s'appuie sur la notion d'arbre de coupes (*componant-tree*) (Salembier et al., 1998). Elle consiste à modéliser explicitement une image à niveaux de gris sous sa forme hiérarchique. Ceci ouvre notamment la voie au développement de techniques de filtrage antiextensif et de segmentation à base d'attributs. De telles approches permettent de traiter efficacement des images dont les niveaux de gris reflètent une sémantique spectrale dans laquelle les objets d'intérêt ont des valeurs extrêmes (maximales ou bien minimales). Au demeurant, elles sont aussi intrinsèquement adaptées au traitement de cartes d'élévation, par exemple obtenues par des données de type LiDAR (Mongus et Zalik, 2014). Suivant le même concept, une version de hiérarchie auto-duale a été proposée, sous la forme d'un arbre de formes (*tree*

of shapes) (Monasse et Guichard, 2000), qui modélise cette fois les relations d'inclusion entre les composantes de l'image, au même titre que le feraient des isohypses dans une carte topographique.

Dans la même veine « topographique », la seconde approche hiérarchique proposée repose sur le concept de ligne de partage des eaux (*watershed*) (Vincent et Soille, 1991). Là encore, l'objectif consiste à considérer une image à niveaux de gris comme un relief topographique. Le but est alors de l'inonder progressivement pour faire apparaître des lignes de partage des eaux sur les crêtes, aboutissant à une partition de l'image. Cette approche est généralement appliquée sur une image préalablement traitée par un opérateur de gradient, visant à mettre en évidence les variations brutales de valeurs de l'image. La première conséquence est que des images multi-spectrales peuvent ainsi être considérées. La seconde conséquence est que cet opérateur de gradient permet de mettre en place une mesure de saillance, qui peut alors servir à établir une hiérarchie de *watersheds* (Najman et Schmitt, 1996), dont le degré de détail est directement lié à la mesure de saillance.

Cette approche, qui permet d'aboutir à différents degrés de segmentation, se trouve à l'interface entre deux notions de hiérarchies : celle directement liée à l'image, et celle liée à ses différentes échelles d'interprétation, discutée ci-après.

2.2. Hiérarchies d'échelles

Le contenu visuel d'une image peut également s'organiser sous la forme d'une hiérarchie donnant lieu à différentes échelles d'interprétation. Afin d'explicitier ce paradigme, considérons l'exemple de la cartographie urbaine, sur lequel nous reviendrons en section 4. Dans ce domaine, les experts utilisent les images pour analyser le territoire à plusieurs niveaux sémantiques, chacun correspondant à une échelle spatiale de représentation, dans le but d'extraire différents types d'objets d'intérêt : quartiers, blocs urbains, objets urbains individuels, etc.

Les différents objets d'intérêt relatifs à ces niveaux d'analyse sont généralement organisés hiérarchiquement : un quartier se décompose en un ensemble de blocs urbains, pouvant eux-mêmes se décomposer en ensembles d'objets urbains « simples » (toitures, arbres, routes, etc.). Au demeurant, du fait de la diminution de la taille des pixels ou de l'augmentation de la résolution spatiale des données, ces derniers objets peuvent eux-mêmes se décomposer, de manière hétérogène, en éléments dont la sémantique devient moins évidente (par exemple, les matériaux d'un même bâtiment).

Pour extraire et analyser ces différents niveaux d'objets d'intérêt, de nombreuses stratégies multi-échelles ont été proposées. Historiquement, les premières approches – essentiellement linéaires – se sont appuyées sur des espaces d'échelles obtenus par convolutions (Lindeberg, 2004), ou dans un cadre d'analyse fréquentielle. Alternativement, des approches plus directement liées à l'espace support des images ont également été proposées. Dans ce contexte, les pyramides régulières

(par exemple, les *quadrees* (Wang, 1991)), prenant mal en compte la géométrie des objets d'intérêt, ont progressivement laissé la place à des pyramides non-régulières (Montanvert et al., 1991).

Cette non-régularité se prête particulièrement à l'utilisation d'approches non-linéaires, notamment par le biais de filtres morphologiques (Benediktsson et al., 2003; Gueguen et al., 2010). à titre d'exemple, des compositions d'ouvertures et de fermetures morphologiques basées sur des éléments structurants de taille croissante permettent de générer des profils morphologiques en chaque pixel de l'image, aboutissant à leur caractérisation multi-échelle (Pesaresi et Benediktsson, 2001). Une limitation de telles approches réside néanmoins dans leur incapacité à permettre la mise en correspondance immédiate entre les différents niveaux de hiérarchie obtenus. En d'autres termes, un segment extrait d'un niveau particulier ne se décomposera généralement pas en sous-segments aux niveaux suivants.

Afin de pallier ce problème, il est possible de s'appuyer sur des approches non plus basées sur les pixels, mais sur des régions de l'image. Dans ce paradigme, une région à un niveau de hiérarchie donnée doit – par construction – s'exprimer comme l'union de régions aux niveaux suivants. Ces hiérarchies peuvent être construites de deux manières opposées. Ainsi, dans les approches « descendantes » (*top-down*), les partitions les plus grossières sont d'abord produites, puis raffinées en partitions de plus en plus fines, par des processus spécifiques de décomposition (Shi et Malik, 2000). Toutefois, les approches les plus utilisées sont celles reposant sur un principe « ascendant » (*bottom-up*). Dans ce cas, les partitions les plus fines sont d'abord produites. Les segments adjacents les plus similaires (relativement à des propriétés particulières (Baatz et Schape, 2000; Akcay et Aksoy, 2008)) sont ensuite agrégés pour former des segments plus larges pouvant correspondre à des objets d'intérêt plus complexes.

Un exemple classique d'approche ascendante, développé dans le cadre de la morphologie mathématique, est celui des arbres binaires de partitions (*binary partition trees* ou BPT) (Salembier et Garrido, 2000). Les feuilles des arbres correspondent alors aux segments basiques de l'image – en général ses zones plates – tandis que les nœuds suivants correspondent aux segments issus des fusions de ces régions élémentaires. La priorité de ces fusions est définie par une mesure de similarité entre les segments. En conséquence, un BPT représente une modélisation de l'image à différentes échelles, et donc entre différentes familles d'objets d'intérêt, au regard de cette mesure (Vilaplana et al., 2008). Nous reviendrons plus en détails sur la notion d'arbre binaire de partitions dans la section 3.1.

La capacité à pouvoir représenter une image donnée à différents niveaux d'échelles, grâce à des structures hiérarchiques modélisant les différentes familles d'objets d'intérêt, permet de développer des stratégies de traitement – et notamment de segmentation – efficaces dans de nombreux cadres applicatifs. Toutefois, la validité de

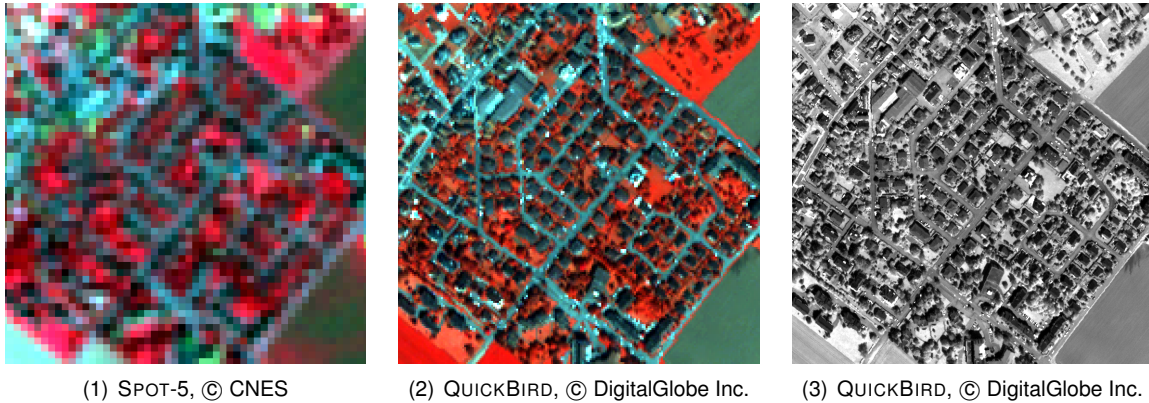


FIGURE 1: Images satellitaires d'une même zone géographique à différentes résolutions spatiales. (1) Résolution moyenne (1 pixel = 9.6 m × 9.6 m), quatre bandes spectrales. (2) Haute résolution (1 pixel = 2.4 m × 2.4 m), quatre bandes spectrales. (3) Très haute résolution (1 pixel = 60 cm × 60 cm), une bande spectrale.

ces approches trouve ses limites lorsque la corrélation entre l'homogénéité sémantique et l'homogénéité spectrale de l'image diminue. Afin de répondre à cette nouvelle question, il est possible de prendre en compte – toujours dans un paradigme hiérarchique – non plus seulement une image, mais une famille d'images d'une même scène à plusieurs résolutions.

2.3. Hiérarchies de résolutions

La dernière décennie a été marquée par une très forte augmentation du nombre de capteurs d'images disponibles, permettant de diversifier les sources de collecte des données, et de s'adapter aux besoins des utilisateurs. En particulier, la multiplication des missions spatiales/aériennes permet de recueillir une quantité importante de données multi-résolutions (Figure 1). Grâce au géoréférencement, il est alors possible de mettre en correspondance les acquisitions disponibles d'une même scène, conduisant à des hiérarchies de résolutions. Ces hiérarchies représentent des ensembles de données complémentaires et fournissent des représentations différentes des territoires observés.

Pour pouvoir analyser ces données à différents niveaux sémantiques, il est – comme dans le cas des hiérarchies d'échelles – nécessaire d'extraire différents niveaux d'objets d'intérêt. Toutefois, les approches d'analyse multi-échelles évoquées précédemment ne peuvent plus – du fait de leur nature mono-image – être employées directement ici. Face à ce problème, des approches d'analyse dédiées au traitement de données multi-résolutions ont été proposées. On peut notamment distinguer deux familles : celle procédant par fusion des différentes données disponibles, et celle analysant conjointement les différentes images.

Les approches par fusion agrègent les données afin de construire une vue unique (*pansharpening*) (Witharana et al., 2013) sur laquelle est alors appliquée une analyse hiérarchique afin d'extraire les classes d'objets recherchées. Ces approches sont principalement utilisées pour l'analyse d'images multi-résolutions acquises

par des satellites produisant simultanément plusieurs images de la même zone : une image panchromatique et une image multi-spectrale. La première présente une haute résolution spatiale mais une faible résolution spectrale, tandis que la seconde présente une haute résolution spectrale mais une faible résolution spatiale. De nombreuses méthodes ont été envisagées pour produire une image finale de bonne qualité spatiale et spectrale (Dou et al., 2007). Néanmoins, des pertes d'informations induites par la fusion des données sont inévitables.

Les approches analysant conjointement les différentes images sans les fusionner sont généralement basées sur des processus collaboratifs échangeant des informations entre les différentes résolutions. Ces processus permettent d'utiliser toutes les informations disponibles de manière conjointe afin de tirer parti de la complémentarité des données. De nombreux travaux ont déjà été réalisés sur la collaboration de méthodes de classification non-supervisée pour le cas multi-source (Bruzzone et al., 2002; Wemmert et al., 2009). Ces méthodes font collaborer plusieurs stratégies non-supervisées pour tenter de trouver un consensus sur la classification des images. Néanmoins, elles travaillent généralement sur les pixels de l'image : ils leur est donc difficile d'extraire différents niveaux d'objets d'intérêt.

Afin d'étendre ce paradigme collaboratif au niveau des régions des images, il est nécessaire de pouvoir mettre en correspondance des régions d'intérêt à travers la hiérarchie de résolutions. à l'instar des hiérarchies multi-échelles, deux stratégies sont généralement considérées pour construire de telles hiérarchies multi-résolutions : les approches ascendantes et les approches descendantes. Les approches ascendantes (Gaetano et al., 2009) opèrent préalablement par une segmentation de l'image de plus haute résolution, aboutissant généralement à une sur-segmentation, dans le but de préserver les détails géométriques les plus élémentaires. Les segments de cette sur-segmentation initiale sont ensuite fusionnés hiérarchiquement, en utilisant les informations radiométriques et géométriques

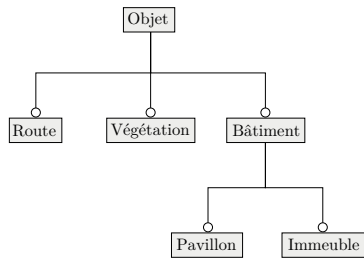


FIGURE 2: Exemple de hiérarchie de six concepts urbains.

des images aux résolutions décroissantes. Dans ce contexte, les informations les plus prépondérantes sont celles obtenues aux résolutions les plus hautes. Afin de mieux équilibrer l'usage de la globalité des informations disponibles, la stratégie descendante vise quant à elle à analyser l'image à la résolution la plus grossière puis à l'affiner progressivement par exploration des résolutions supérieures. Il est alors possible d'extraire et de classifier des segments correspondant à des objets complexes, tout en s'abstrayant du bruit sémantique induit par les détails (Goffe et al., 2011). Cette seconde stratégie est comparable au principe de la vision humaine (Mallat, 1996), et a notamment été considérée pour l'analyse conjointe d'images HR et MR (Sun et al., 2003).

Les besoins actuels en analyse d'images multi-résolutions justifient l'emploi de processus collaboratifs segmentation-classification permettant d'obtenir une analyse multi-échelle d'une scène à partir de hiérarchies de résolutions. Ce point sera illustré en section 3.2.

2.4. Hiérarchies de concepts

Face à l'explosion de la complexité des données, les approches classiques (non-supervisées) de segmentation et de classification d'images atteignent leurs limites et peuvent notamment conduire à des résultats insuffisants face aux besoins applicatifs. L'augmentation du nombre de données à traiter, du nombre d'attributs à définir, ou encore du nombre de classes à extraire, sont des exemples des problèmes les plus couramment rencontrés. Les nouvelles tendances en analyse d'images de télédétection (Kopanas et al., 2002) visent alors à s'appuyer sur l'intégration de connaissances *a priori* de haut-niveau, relatives au contenu des images.

Les connaissances du domaine sont des informations géographiques que l'expert introduit dans le processus d'analyse. Ces informations sont relatives à un certain type de territoire, et ne sont donc pas nécessairement spécifiques aux jeux de données considérés. Plusieurs types d'informations sont envisageables : des connaissances mono-échelles (par exemple relatives à l'occupation des sols observés à partir d'une image à une échelle spécifique), ou bien multi-échelles (par exemple relatives à l'organisation spatiale et à la composition des objets d'intérêt à différents niveaux d'échelles).

Ces connaissances peuvent elles-mêmes être organisées sous la forme de hiérarchies de concepts, connues également sous le nom d'ontologies. Une ontologie (Gruber, 1995) est un réseau sémantique qui re-

groupe un ensemble de concepts décrivant un domaine. Ces concepts sont liés les uns aux autres par des relations taxonomiques d'une part, et sémantiques d'autre part. Les ontologies fournissent ainsi un cadre formel de haut niveau pour la représentation de connaissances. Un exemple de hiérarchie de concepts est présenté dans la figure 2. En haut se situe le concept le plus général *Objet*. Dans cet exemple, trois concepts le spécialisent : *Bâtiment*, *Route* et *Végétation*. Le concept *Bâtiment* peut lui-même se spécialiser en *Immeuble* et *Pavillon*.

Pour pouvoir utiliser les connaissances ainsi formalisées, il est indispensable de les compléter afin de leur faire correspondre une réalité physique dans l'image. à cette fin, plusieurs types de caractéristiques peuvent être utilisés : intervalles sur des attributs quantitatifs, méta-connaissances sur les attributs discriminants du concept, attributs qualitatifs, etc. Les concepts présents dans la hiérarchie peuvent alors être utilisés pour procéder à la segmentation, l'identification et/ou la classification des objets d'intérêt présents dans l'image.

Ces hiérarchies de connaissances ont notamment été employées dans des approches de segmentation multi-échelle afin de guider les processus de fusion de régions pour la création de hiérarchies de segments. à titre d'exemple, dans (Inglada et Michel, 2009), une approche par graphes est proposée, où l'image est décrite par un ensemble de segments primitifs reliés entre eux par des relations spatiales. Une fois ces relations établies, il est possible de définir des règles expertes de fusion conduisant à la formation de régions correspondant à des objets complexes, puis de rechercher dans le graphe l'ensemble des sous-graphes similaires à ces exemples. Ainsi le concept *Pavillon* peut être décrit dans une ontologie par des relations de voisinage fréquentes : un pavillon est situé à côté d'une parcelle de végétation et d'une route. De telles règles peuvent ultérieurement être utilisées pour inférer des segments correspondant à des objets composites à partir de segments primitifs.

Les hiérarchies de concepts peuvent également être employées pour la caractérisation sémantique de régions préalablement obtenus *via* une étape de segmentation. à titre d'exemple, dans (Forestier et al., 2012), un score d'appariement est proposé, s'appuyant sur la similarité entre les valeurs des attributs d'une région et l'espace de définition de cet attribut pour un concept. Ce score est composé d'une similarité locale (la correspondance d'une région avec un concept donné) et d'une similarité globale (la correspondance d'une région avec une hiérarchie de concepts). Des recherches sont également en cours afin de permettre d'affecter à ces concepts des connaissances relatives aux relations spatiales. Ainsi, le concept *Pavillon* pourrait être caractérisé par la propriété : « souvent adjacent à une parcelle de végétation » ; il pourrait aussi s'exprimer en termes de composition, notamment par son appartenance au concept *Bloc urbain d'habitations individuelles*, afin de relier ces concepts dans le but de mieux les caractériser.

Nous reviendrons plus en détails sur la notion de hiérarchies de concepts dans la section 3.3.

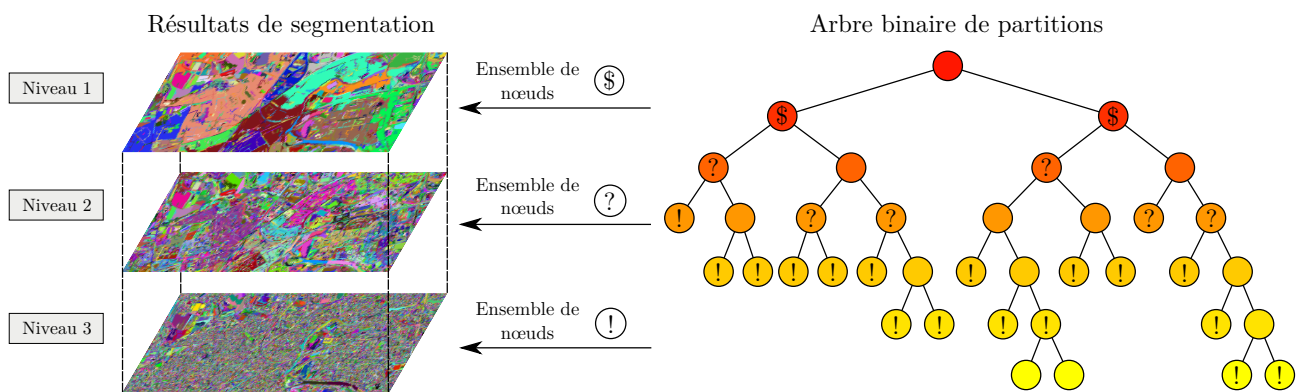


FIGURE 3: Un exemple d'arbre binaire de partitions associé à une image THR (pour des raisons de lisibilité, le nombre de nœuds a été artificiellement réduit). Trois niveaux de segmentation (associées à trois coupes de l'arbre) sont représentés.

3. Un exemple d'approche multi-hiérarchique

Nous présentons ci-après une approche multi-hiérarchique pour l'analyse d'images satellitaires combinant les différents paradigmes présentés précédemment. Cette approche est composée de trois briques méthodologiques pouvant être considérées de manière individuelle ou composées les unes avec les autres pour l'analyse multi-échelle d'images multi-résolutions.

La première brique (section 3.1) est dédiée à la segmentation hiérarchique d'une image et repose sur une méthode par arbres binaires de partitions.

La deuxième brique (section 3.2) est dédiée à l'analyse multi-échelle d'un ensemble d'images multi-résolutions et repose sur une méthode impliquant des phases de segmentation et de classification.

Finalement, la troisième brique (section 3.3) est dédiée à la classification de régions d'intérêt, extraites d'images multi-résolutions par une étape de segmentation, et repose sur une distance hiérarchique basée sur la notion de hiérarchie de concepts.

3.1. Analyse hiérarchique multi-échelle : segmentation par arbres binaires de partitions

Dans le cadre de travaux récents (Kurtz et al., 2011a,c) nous avons proposé une approche de segmentation par arbres binaires de partitions (BPT) combinant phases interactives et automatiques. Elle repose sur la définition d'exemples de segmentation sur des parties spécifiques de l'image, automatiquement reproduits sur le reste des données. Cette interaction permet d'impliquer l'expert au cœur du système afin de pallier certains problèmes liés à la définition automatique d'une coupe dans l'arbre pour la segmentation d'images (T)HR.

Construction du BPT. Le BPT d'une image \mathcal{I} est construit de manière ascendante, c'est-à-dire des feuilles vers la racine. Partant d'une segmentation initiale du support de l'image E , généralement composée des zones plates de \mathcal{I} , les nœuds de l'arbre sont définis successivement par la fusion de couples de segments (nœuds) voisins et similaires dans l'image.

De nombreux BPT peuvent ainsi être obtenus pour une même image. Afin de décider quel arbre est le plus pertinent, il est nécessaire de définir une priorité de fusion parmi les nœuds. La construction d'un BPT conduit ainsi à deux notions principales : le *modèle de région*, qui spécifie comment les segments/nœuds sont caractérisés ; et le *critère de fusion*, qui définit la similarité entre les nœuds voisins ainsi que leur ordre de fusion. Les modèles et critères de base utilisés dans la plupart des approches de segmentation par fusion de segments, sont généralement fondés sur l'homogénéité radiométrique des nœuds (Garrido et al., 1998; Felzenszwalb et Huttenlocher, 2004). Dans le cadre de nos travaux, nous nous sommes également appuyés sur des propriétés géométriques, tels que des critères d'élongation (routes, voies ferrées, canaux, etc., en zones urbaines) ou encore des critères géomorphologiques (pentes, altitudes issues de données LiDAR, en zones montagneuses). Un BPT spécifique peut ainsi être produit pour une image en utilisant un modèle de région et un critère de fusion donnés. La figure 3 illustre un exemple de construction d'un BPT.

Segmentation à partir d'un BPT. Une fois le BPT construit, il convient de définir une coupe au sein de l'arbre pour obtenir une partition de l'image et ainsi un découpage des objets d'intérêt de la scène.

Dans le cadre de l'analyse d'images (T)HR, différents objets peuvent être extraits à différentes échelles. En effet, les objets d'intérêt relatifs à un même niveau d'analyse peuvent apparaître comme plus ou moins complexes et hétérogènes au sein d'une même image. Par exemple, les zones urbanisées vont apparaître plus hétérogènes que les zones rurales ou les forêts alors que ces objets font partie d'un même niveau d'analyse.

Dans la plupart des cas, une segmentation unique ne permet pas d'extraire des segments correspondant à tous les objets d'intérêt relatifs à un même niveau d'analyse. Lors de l'utilisation d'une segmentation hiérarchique, ces objets peuvent ainsi émerger à différents niveaux de la hiérarchie obtenue. Il apparaît alors difficile de trouver une coupe unique dans un BPT, permettant de les extraire simultanément.

Une solution généralement employée consiste à utiliser plusieurs niveaux de la hiérarchie de segmentations

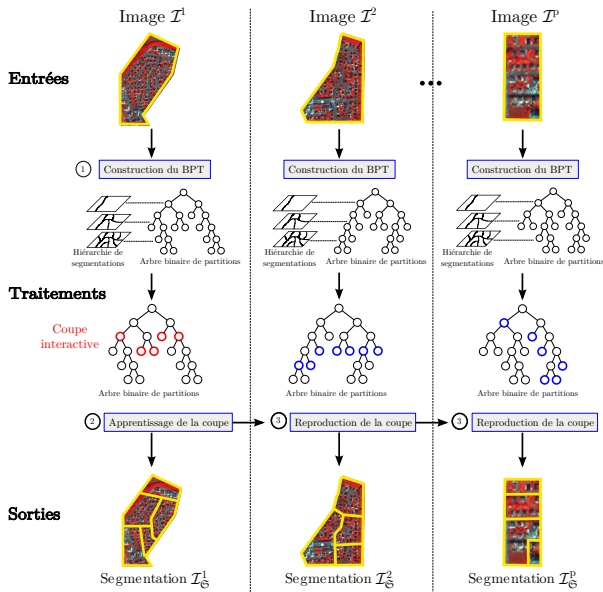


FIGURE 4: Approche de segmentation hiérarchique par BPT. En rouge : interactions de l'utilisateur. En bleu : traitements automatiques. En jaune : p imagettes représentant p zones différentes mais centrées sur le même type d'objets d'intérêt.

afin d'extraire au mieux les différents objets recherchés. Cependant, cette solution est consommatrice de temps et peut engendrer des erreurs. Nous proposons donc d'adapter la notion de segmentation par BPT pour pallier ce problème et aider l'expert dans cette étape de segmentation multi-niveau.

Approche interactive basée sur des exemples. Pour prendre en compte ce phénomène d'échelle dans le processus de segmentation hiérarchique, nous le divisons en un ensemble de sous-processus, chacun adapté à l'extraction d'une catégorie d'objets d'intérêt. Au lieu de segmenter l'intégralité d'une image en utilisant un seul niveau de la hiérarchie du BPT, l'idée est d'adapter l'approche de segmentation à des zones locales de l'image représentant le même type d'objets, de même homogénéité radiométrique. Une telle approche permet de segmenter localement les différents objets d'intérêt de l'image, tout en évitant les problèmes liés à la superposition des résultats de segmentation.

Le principe de cette approche est de construire plusieurs BPT, chacun sur une zone particulière de l'image. Cette image peut alors être divisée en un ensemble de q imagettes, centrées sur les différents objets d'intérêt que l'expert souhaite extraire de la scène. Une fois les q arbres associés construits, on extrait une coupe spécifique par BPT, dans le but d'obtenir des niveaux de segmentation différents pour chacune des q imagettes. Comme énoncé précédemment, nous proposons d'utiliser une approche interactive pour choisir une coupe pertinente. Une telle approche implique l'intervention de l'expert pour chacun des q BPT résultants.

Pour minimiser cette tâche, nous utilisons un mécanisme interactif de définition d'exemples de coupes, mais aussi un mécanisme automatique de reproduction de

ces exemples dans le reste de l'image. Cette approche de segmentation prend en entrée p imagettes, représentant p zones différentes mais centrées sur le même type d'objets d'intérêt, extraites des q zones définies précédemment. Elle retourne p segmentations (Figure 4). Pour l'une de ces imagettes, un BPT est construit, puis l'utilisateur définit interactivement une coupe, produisant ainsi une segmentation. Cette segmentation caractérise le niveau d'échelle souhaité par l'utilisateur et est utilisée comme exemple pour segmenter les $p - 1$ imagettes restantes. L'algorithme d'apprentissage de l'exemple de coupe est détaillé dans (Kurtz et al., 2011a).

Il est possible d'appliquer plusieurs fois cette approche dans une même image, pour segmenter différents types d'objets ayant des sémantiques et des niveaux d'échelle différents, et traiter ainsi l'ensemble des q imagettes définies précédemment. En conséquence, toutes les imagettes centrées sur le même type d'objets d'intérêt sont segmentées d'une manière similaire.

3.2. Analyse hiérarchique multi-résolution : segmentation et classification croisées

Dans (Kurtz et al., 2012a) nous avons proposé une approche d'analyse hiérarchique d'ensembles d'images multi-résolutions permettant l'étude d'un territoire à différents niveaux d'échelle. Cette approche repose sur une stratégie impliquant des phases de segmentation (réalisées *via* l'approche par arbres binaires de partitions défini ci-avant) et de classification croisées appliquées de manière descendante à travers la résolution.

Segmentation-classification croisée. Pour mettre en œuvre cette stratégie, nous faisons collaborer une approche de segmentation et une approche de classification non-supervisée afin de construire directement des hiérarchies de segments d'une manière descendante à travers la résolution. Le principe consiste à extraire des segments grossiers des images aux plus faibles résolutions (pour analyser les grandes zones de la scène), puis à les affiner successivement et itérativement en utilisant les images aux résolutions les plus fines (pour analyser des structures plus élémentaires).

L'utilisation d'une approche multi-résolution descendante permet à l'utilisateur de restreindre la quantité de données traitées (par un « zoom » progressif) et ainsi de diminuer les problèmes liés à la complexité mémoire des données analysées. Cette approche permet également de construire et de classifier des segments, à travers la résolution, pouvant être mis en correspondance d'une image à une autre. Une telle propriété rend possible, par la suite, la mise en place d'une collaboration entre les régions issues des différentes images.

Analyse descendante d'images multi-résolutions. L'approche proposée requiert en entrée :

- un ensemble multi-résolution $\{\mathcal{I}^t\}$ de n images d'une même scène, ayant des résolutions spatiales croissantes $\tau_1 \leq \dots \leq \tau_n$, et pouvant avoir été produites par des capteurs différents (et posséder ainsi des bandes spectrales différentes) ;

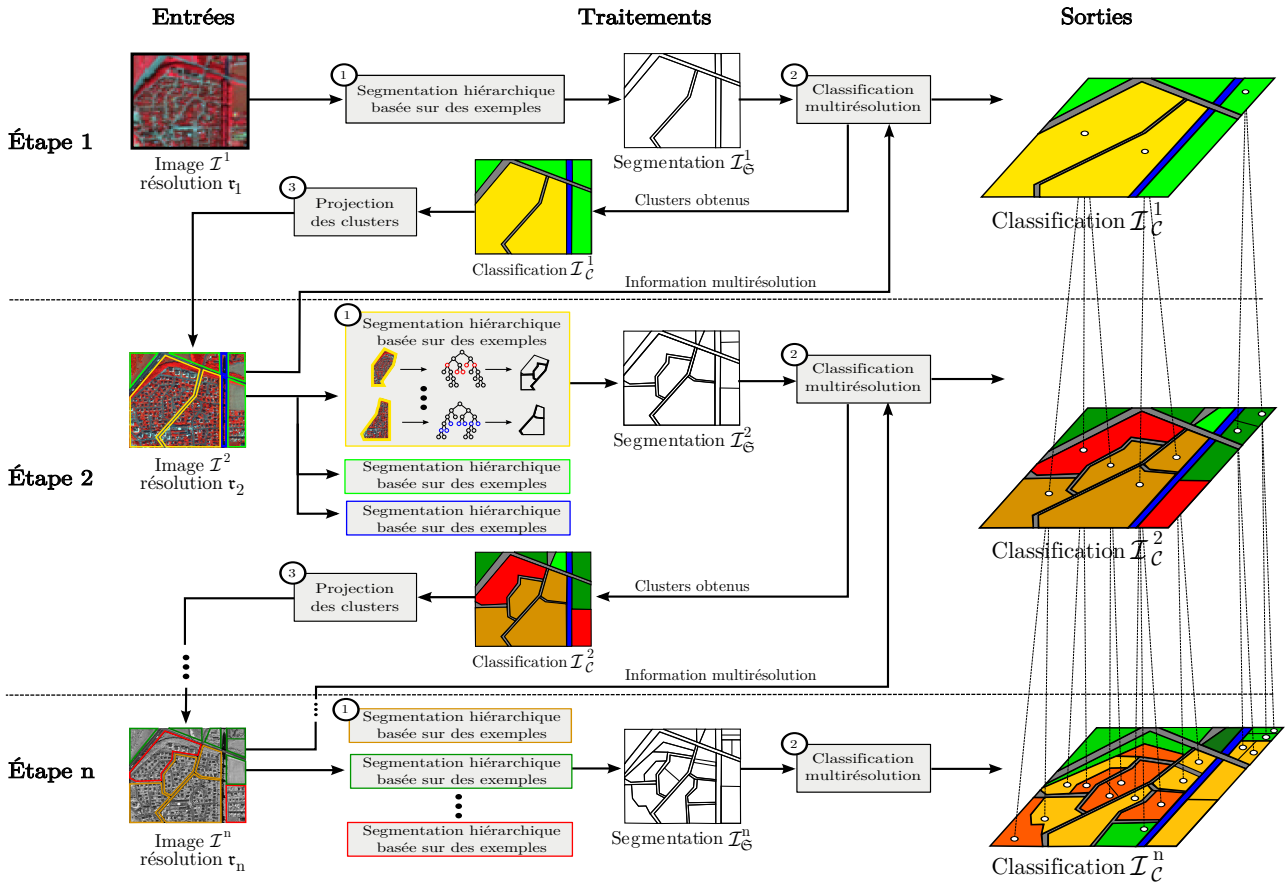


FIGURE 5: Illustration de l'approche de segmentation-classification croisée pour l'analyse d'images multi-résolutions.

et fournit en sortie :

- un ensemble $\{\mathcal{I}_S^t\}$ de n images segmentées, associé à un ensemble $\{\mathcal{I}_C^t\}$ de n images classifiées (une par résolution considérée), organisées hiérarchiquement, permettant différents niveaux d'interprétation.

Cette approche fonctionne en n étapes successives, chaque étape étant réalisée dans le but d'analyser une image \mathcal{I}^t parmi les n images (en partant de l'image \mathcal{I}^1 à la plus faible résolution, jusqu'à l'image \mathcal{I}^n à la résolution la plus élevée). Finalement, n niveaux de segmentation et de classification sont donc produits.

Ce processus étant basé sur le paradigme régions, chacune des n étapes est composée de :

1. une approche de segmentation hiérarchique (basée sur des exemples, cf. section 3.1) ;
2. une approche de classification multi-résolution (décrite ci-après) ;
3. une projection des clusters obtenus dans la résolution suivante.

Pour traiter des ensembles d'images multi-résolutions, ces trois sous-étapes sont appliquées successivement à travers la résolution (Figure 5). À chaque étape t , le résultat de l'étape précédente $t - 1$ (un ensemble de régions labellisées en différentes catégories) est projeté dans l'image à la résolution courante \mathcal{I}^t et traité comme

entrée de la méthode. Le principe est alors de décomposer (c'est-à-dire de segmenter puis de classier) ces différentes catégories de régions dans la résolution courante afin de les affiner. Pour ce faire, à chaque étape t , l'approche de segmentation par BPT est appliquée afin de partitionner l'ensemble des catégories de régions fournies par l'étape $t - 1$. Une fois que toutes les régions composant ces différentes catégories ont été décomposées, il est possible de créer une partition globale de l'image \mathcal{I}^t en regroupant tous les segments résultant de cette étape. Ces segments sont ensuite caractérisés par des attributs multi-résolutions, puis sont classifiés en différents ensembles homogènes par l'approche de classification multi-résolution non-supervisée présentée ci-après, formant ainsi de nouvelles catégories de régions autorisant un niveau d'analyse différent. Ce résultat est alors projeté dans l'image \mathcal{I}^{t+1} , de résolution plus fine, pour que le processus y soit ré-appliqué.

Classification de segments multi-résolutions. L'étape de classification employée dans l'approche d'analyse descendante décrite ci-dessus repose sur une méthode de classification des segments caractérisés par des attributs multi-résolutions (Kurtz et al., 2011b).

Dans les approches basées régions classiques, les segments à classier (extraits d'une image à une résolution r_t) sont généralement décrits par des attributs radiométriques ou géométriques calculés dans l'image à

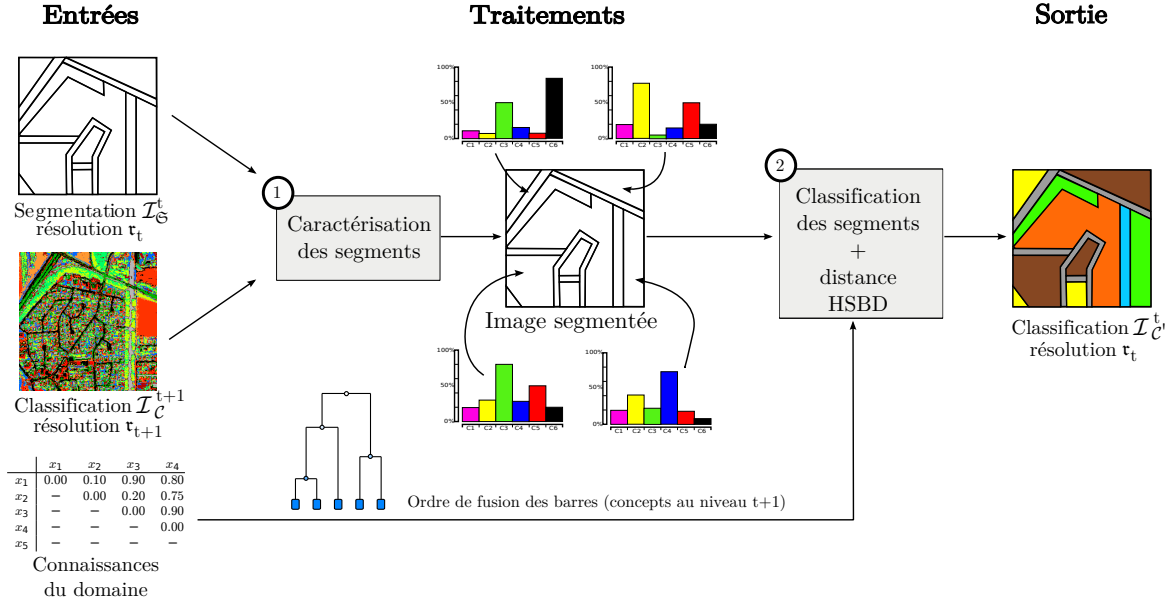


FIGURE 8: Illustration du processus de post-classification mono-résolution appliqué au traitement d'un niveau de résolution.

résultats de classification. Ce dernier est basé sur une post-classification des régions obtenues par l'approche multi-résolution descendante via l'approche de classification multi-résolution proposée dans la partie précédente, combinée à la distance HSBD.

HSBD : une distance hiérarchique basée sur la sémantique pour la classification. La distance hiérarchique (*Hierarchical Semantic-Based Distance*) permet la comparaison d'histogrammes de composition liés à une sémantique (Kurtz et al., 2013).

Le calcul de cette distance entre deux histogrammes $H(A)$ et $H(B)$, requiert deux paramètres : une matrice de dissimilarité \mathcal{M}^{dis} modélisant les valeurs de proximité sémantique entre les instances de $H(A)$ et $H(B)$ et une distance d'histogrammes classique D_{bin} .

Avant de pouvoir calculer la distance HSBD, la stratégie adoptée (basée sur un modèle de calculs fin-à-grossier) nécessite de définir un moyen de fusionner hiérarchiquement les différentes instances représentées par les histogrammes en « clusters » d'instances (c'est-à-dire des instances de niveaux sémantiques plus élevés). Cette étape de pré-traitement repose principalement sur la construction d'un dendrogramme \mathcal{D} induit par \mathcal{M}^{dis} et modélisant la hiérarchie de fusion des instances. Un exemple de dendrogramme modélisant la similarité sémantique entre des objets urbains individuels est présenté par la figure 7.

Une fois le dendrogramme \mathcal{D} construit, la distance HSBD peut être calculée en deux étapes principales :

- **Étape 1. Calcul des sous-distances barre-à-barre hiérarchiques** Durant un processus de fusion itératif scannant chaque étage du dendrogramme (de ses feuilles jusqu'à sa racine), les histogrammes liés à $H(A)$ and $H(B)$, et induits par la fusion des instances composant chaque

cluster de l'étage k courant, sont construits. Après chaque itération, la distance D_{bin} est calculée entre chaque couple d'histogrammes créé. Cette distance « intermédiaire » est notée \mathbf{d}^k .

- **Étape 2. Fusion des sous-distances barre-à-barre** Les distances \mathbf{d}^k calculées pour tous les étages du dendrogramme, et l'énergie sémantique nécessaire pour aller d'un étage à l'autre (correspondant à la hauteur $h_{\mathcal{D}}$ du dendrogramme entre deux étages), sont ensuite fusionnées en une fonction qui est finalement intégrée pour fournir la valeur de la distance HSBD.

Nous définissons ainsi la distance HSBD entre $H(A)$ et $H(B)$ comme une intégrale finie égale à :

$$\frac{1}{2} \sum_{k=0}^{s-2} \left[(\mathbf{d}^{k+1} + \mathbf{d}^k) (h_{\mathcal{D}}(k+1) - h_{\mathcal{D}}(k)) \right] \quad (1)$$

Intégration de la distance HSBD dans l'approche de classification multi-résolution. Pour comparer et classifier des régions décrites par des histogrammes de composition, nous intégrons la distance HSBD dans l'algorithme de classification utilisé dans l'approche de classification multi-résolution (Figure 8). Pour comparer des régions extraites d'une image à une résolution τ_t avec HSBD, l'utilisateur doit fournir une matrice de dissimilarité \mathcal{M}^{dis} qui modélise les relations sémantiques entre chaque instance du concept sémantique relatif aux clusters extraits de l'image à la résolution τ_{t+1} .

L'intégration de la distance HSBD au sein de l'approche de classification multi-résolution permet ainsi de prendre en compte des connaissances de haut niveau (apportées par l'utilisateur à une résolution τ_{t+1}), pour améliorer la comparaison des données et ainsi les résultats de classification obtenus à une résolution τ_t .

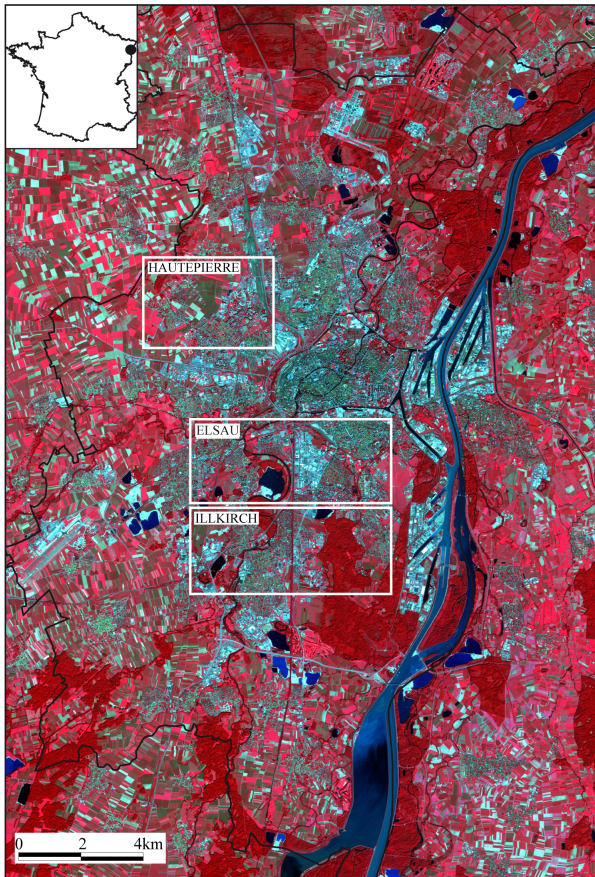


FIGURE 9: Illustration de la zone urbaine de Strasbourg, France (SPOT-5, © CNES – 2002). Les sites d'étude (ELSAU, ILLKIRCH, HAUTEPIERRE) sont représentés en blanc sur une image à moyenne résolution (10 mètres).

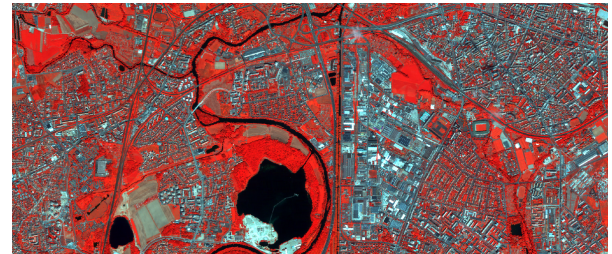
4. Un cas d'étude : analyse des milieux urbains

Nous présentons maintenant un cas d'application de l'approche multi-hiérarchique décrite en section 3. Sa finalité est d'extraire des informations fines dans des images à très haute résolution. Néanmoins, l'étude de ces seules images n'est pas suffisante pour obtenir des résultats satisfaisants. En effet, leur taille élevée rend leur analyse délicate du fait d'une grande masse d'information qui masque les données pertinentes, et d'un degré de détail très fin qui, de prime abord, nuit à la pertinence de l'analyse.

Dans ce contexte, une approche incrémentale – et hiérarchique – est justifiée, non seulement pour pouvoir progressivement se focaliser sur les éléments sémantiques d'intérêt, mais aussi pour pouvoir intégrer, au niveau de détail adéquat, des éléments d'information susceptibles de guider l'analyse. Une description plus complète de cette application pourra être trouvée dans (Kurtz et al., 2012a). Un cas d'étude alternatif, portant pour sa part sur des images de glissements de terrain, est également décrit dans (Kurtz et al., 2014). On pourra noter, qu'entre ces deux applications, les informations *a priori* intégrées dans le schéma méthodologique diffèrent, et sont notamment issues des besoins thématiques.



(1) SPOT-5, © CNES – 2002



(2) QUICKBIRD, © DigitalGlobe – 2002



(3) QUICKBIRD, © DigitalGlobe – 2002

FIGURE 10: Présentation du jeu de données ELSAU (6 576 m × 2 793 m). (1) Résolution moyenne (1 pixel = 9.6 m × 9.6 m), quatre bandes spectrales. (2) Haute résolution (1 pixel = 2.4 m × 2.4 m), quatre bandes spectrales. (3) Très haute résolution (1 pixel = 60 cm × 60 cm), une bande spectrale.

4.1. Contexte

La cartographie des milieux urbains représente un domaine d'activité crucial. En effet, les collectivités chargées de la gestion et de la planification du développement urbain, ou encore les agences de gestion de crises environnementales et de catastrophes naturelles doivent pouvoir gérer et suivre l'évolution de l'urbanisation de façon précise, et en particulier sur les zones à hauts risques pour les populations. Pour ce faire, il est nécessaire de cartographier les zones urbaines, pour obtenir des informations utiles, pouvant être directement exploitables par ces organisations (par exemple pour quantifier le développement urbain, obtenir des cartes des dommages liés à une catastrophe naturelle, etc.).

Dans ce contexte, le recours à des images à très haute résolution a d'ores et déjà démontré un réel intérêt (Puissant et Weber, 2002). De telles images permettent notamment d'avoir accès à des objets urbains (maisons individuelles, jardins, rues, etc.). L'analyse de ces objets est néanmoins complexe, du fait de leur signature spectrale composite et/ou variable. Au demeurant, la nomenclature des objets urbains s'intègre hié-

Niveau échelle	1. Quartiers urbains 1:100 000 ↔ 1:25 000	2. Blocs urbains 1:10 000	3. Objets urbains simples 1:5 000
Objets d'intérêt	* Quartiers à forte densité	* Blocs urbains continus	* Bâtiments/toits - Toits en tuiles rouges - Toits en tuiles grises - Toits réfléchissants
	* Quartiers à faible densité	* Blocs urbains discontinus - Blocs d'habitations individuelles - Blocs d'habitations collectives	* Végétation - Végétation chlorophyllienne - Végétation non-chlorophyllienne
	* Zones industrielles	* Blocs urbains industriels	* Voies de communication - Rues - Parkings
	* Zones boisées	* Zones de végétation urbaine	* Surfaces d'eau - Rivières - étangs
	* Zones agricoles	* Forêts	* Sols nus
	* Surfaces d'eau	* Zones agricoles	* Ombres portées
	* Sols nus	* Surfaces d'eau	
		* Routes	

TABLE 1: Typologies et niveaux d'analyse utilisés pour les territoires urbains à différentes échelles.

rarchiquement dans des nomenclatures de plus bas niveau, qui incluent notamment les blocs urbains (patés de maisons, quartiers, etc.) observables à haute résolution, et les zones urbaines (zones d'habitation, zones industrielles, etc.) observables à moyenne résolution.

4.2. Données

Il est donc pertinent d'analyser les milieux urbains en utilisant des jeux de données à diverses résolutions, comme illustré, par exemple, en Figure 1 où trois images à moyenne, haute et très haute résolutions sont fournies.

Dans le cadre de cette étude, les expériences ont été réalisées sur la zone urbaine de Strasbourg, France (Figure 9). Ses différents sites présentent des environnements urbains typiques, composés d'objets d'intérêt urbains classiques (surfaces d'eau, zones de forêt, zones industrielles, blocs urbains d'habitations individuelles et collectives, zones agricoles, etc.). Le site **ELSAU** (6 576 m × 2 793 m), plus spécialement illustré dans cette section, correspond à une zone péri-centrale.

Chaque site est associé à un jeu de données multi-résolution comportant :

- une image SPOT-5 à moyenne résolution, composée de quatre bandes spectrales (© CNES). La résolution spatiale originale de l'image est de 10 mètres, ré-échantillonnée à 9,6 mètres ;
- un couple d'images QUICKBIRD (© DigitalGlobe Inc.), composé d'une image multi-spectrale à haute résolution (2,4 mètres) à quatre bandes spectrales, et d'une image panchromatique à très haute résolution (60 centimètres).

Toutes les images ont été acquises au cours de l'année 2002, à la même saison. La figure 10 présente les images du site **ELSAU**.

Il existe dans la littérature (Collet et al., 2001) de nombreuses typologies d'objets urbains, permettant une cartographie à une échelle du 1:100 000 au 1:5

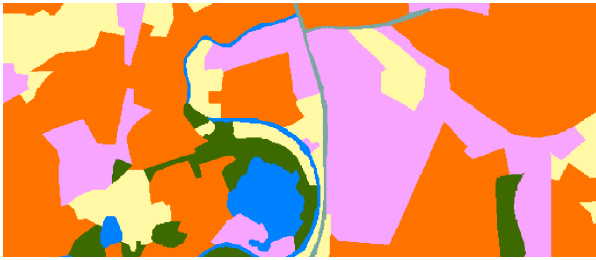
000. Nous présentons ci-après une nomenclature hiérarchique à trois niveaux de sémantique, permettant de cartographier des surfaces urbanisées (Table 1) :

1. le premier niveau est celui des quartiers urbains. Il permet de cartographier le territoire depuis une échelle au 1:100 000 jusqu'à une échelle au 1:25 000, à partir d'images MR, avec un faible niveau de détail. Ce premier niveau de sémantique est employé pour étudier les grandes zones urbaines et leur densité ;
2. le deuxième niveau est celui des blocs urbains, qui permet une analyse du territoire à l'échelle du 1:10 000. Ce niveau a été proposé pour cartographier les blocs délimités par des axes de communication. Il est généralement utilisé pour des images HR ou MR ;
3. le troisième niveau est celui des objets urbains simples. Il permet de cartographier le territoire à l'échelle du 1:5 000 afin d'analyser les objets urbains individuels (bâtiments, routes, etc.) en fonction de leurs matériaux. Il peut être employé pour cartographier le territoire à partir d'images THR.

4.3. Résultats

La méthode multi-hiérarchique décrite en section 3 a donc été appliquée sur ces jeux de données images afin de segmenter, puis de classifier les objets d'intérêt aux divers niveaux d'échelles, associés aux différentes résolutions disponibles.

L'analyse expérimentale d'une telle application requiert de valider divers éléments, parmi lesquels l'efficacité ou la robustesse de la méthode. Dans cet article, nous nous attachons essentiellement à illustrer sa précision, c'est-à-dire l'adéquation entre les résultats obtenus et la réalité. Cette notion de « réalité » s'exprime sous la forme de cartes de vérité-terrain fournies par des experts en cartographie urbaine.



(1) Niveau des quartiers urbains (6 classes sémantiques).



(2) Niveau des blocs urbains (9 classes sémantiques).



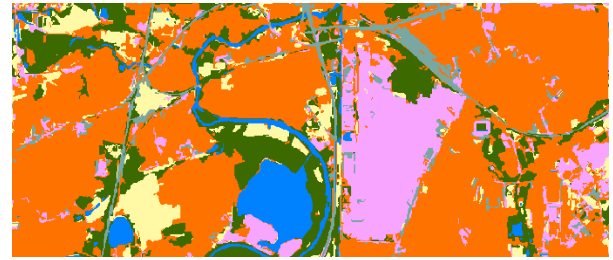
(3) Niveau des objets urbains (1 classe sémantique).

FIGURE 11: Cartes de vérité-terrain associées aux images du jeu de données **ELSAU**.

Un telle vérité-terrain est fournie en figure 11 pour le site **ELSAU**, déjà illustré en figure 10. La première carte contient six classes thématiques correspondant au niveau sémantique des quartiers. Elle a été produite par un expert *via* une interprétation visuelle de l'image MR. La seconde carte contient neuf classes thématiques correspondant au niveau sémantique des blocs urbains. Cette dernière a été extraite d'une base de données régionale (© BDOCS – CIGAL2002²). La dernière carte est utilisée pour évaluer l'extraction des objets urbains simples. Elle contient une unique classe thématique correspondant à l'emprise au sol des bâtiments, et a été extraite d'une base de données produite par l'Institut National de l'Information Géographique et Forestière (BDTOPO © IGN, 2002).

La méthode de segmentation décrite en section 3 a été intégrée dans la bibliothèque logicielle de segmentation JSL, développée en Java tandis que l'approche de classification multi-résolution a été implantée dans la bibliothèque logicielle de clustering JCL, elle aussi développée en Java. Finalement, une interface graphique permettant à l'utilisateur de paramétrer les différents algorithmes liés à ces deux approches (et d'explorer les ré-

2. URL : http://cigalsace.net/recette/produits_cigal.htm.



(1) Classification MR – niveau des quartiers.



(2) Classification HR – niveau des blocs.



(3) Classification THR – niveau des objets urbains.

FIGURE 12: Résultats de classification pour le jeu de données **ELSAU** (6 576 m × 2 793 m). Les couleurs des clusters ont été choisies pour correspondre aux couleurs des classes dans les cartes de vérité-terrain.

sultats), a été développée et intégrée dans la plate-forme Java **MUSTIC** librement téléchargeable³.

Les résultats obtenus ont été évalués quantitativement, par le biais d'indices de qualité (notamment les valeurs de Kappa et de F-mesure), mais aussi qualitativement. Sur ce dernier point, l'illustration fournie en figure 12 permet de cerner les différences entre les résultats obtenus, et ceux attendus (illustrés en figure 11).

Après l'étape de classification, les comparaisons entre les régions classifiées résultantes et les cartes de vérité-terrain permettent d'obtenir des valeurs de Kappa et de F-mesure égales à 0,38 et 0,66 pour le jeu de données **ELSAU**. Les classes thématiques les mieux extraites sont celles correspondant aux quartiers résidentiels, aux quartiers spécifiques, et aux zones agricoles. Cependant, la méthodologie ne permet pas d'extraire directement la classe des routes à cette échelle.

L'étape 2 a été appliquée sur les images HR pour partitionner les quartiers urbains extraits au cours de l'étape précédente, en différents segments pouvant correspondre à des quartiers urbains mixtes, des sous-

3. URL : <http://icube.unistra.fr/fdbt-fr/index.php/Logiciels>

quartiers commerciaux ou industriels, des blocs urbains d'habitations, etc. Les comparaisons entre les régions classifiées résultantes et les cartes de vérité-terrain permettent d'obtenir des valeurs de Kappa et de F-mesure égales à 0,38 et 0,69 pour le jeu de données **ELSAU**. Les classes relatives aux blocs urbains industriels et d'habitations sont particulièrement bien extraites. Les valeurs de F-mesure associées à ces classes sont proches de 0,60 pour tous les jeux de données. On peut cependant remarquer que bien que certains résultats de partition soient composés de différentes régions correspondant correctement aux objets d'intérêt urbains recherchés, ces partitions sont aussi composées de nombreux petits segments représentant les structures linéaires de la scène et les zones de végétation. Ces problèmes de sur-segmentation sont probablement liés aux critères géométriques utilisés par l'algorithme de construction de BPT, qui ne sont pas nécessairement adaptés à l'extraction de zones de végétation.

L'étape 3 a été appliquée sur les images THR pour extraire les objets urbains simples (bâtiments d'habitation collective/individuelle, végétation, rues/parkings, ombres portées, etc.) à partir des blocs urbains extraits précédemment. Du fait de la non disponibilité de toutes les informations de classe pour les cartes de vérité-terrain correspondant aux images THR, nous avons uniquement évalué les précisions relatives aux classes de bâtiments. Les comparaisons entre les régions classifiées résultantes et les cartes de vérité-terrain permettent d'obtenir des valeurs de Kappa et de F-mesure égales à 0,38 et 0,71 pour le jeu de données **ELSAU**. De plus, les comparaisons réalisées visuellement montrent que les autres catégories d'objets d'intérêt urbains simples sont extraits de façon pertinente, bien que les bords des segments ne soient pas parfaits.

5. Conclusion

Les travaux résumés dans cet article constituent des exemples représentatifs des différentes opportunités méthodologiques et applicatives offertes par les approches hiérarchiques, dans le cadre de l'analyse d'images de télédétection. Ces approches constituent au demeurant un domaine de recherche en pleine expansion. Cette rapide évolution induit des avancées non seulement en ce qui concerne les concepts fondamentaux et les enchères méthodologiques, en amont des applications, mais aussi en ce qui concerne l'usage qui en est fait d'un point de vue applicatif. La présente section conclut ainsi cette étude par l'énumération – non exhaustive – de pistes de recherche prometteuses, à plus ou moins long terme.

5.1. Pistes à court terme

Classiquement, les approches hiérarchiques s'appuient sur des entités organisées de manière arborescente. Les structures de données induites permettent ainsi le développement de schémas algorithmiques efficaces, basés sur des paradigmes de type « diviser pour

régner ». Des travaux récents visent toutefois à explorer désormais aussi des hiérarchies qui ne sont plus des arbres, mais peuvent prendre la forme de structures plus complexes, telles que des graphes orientés acycliques. De telles approches sont notamment pertinentes dans le cas de l'analyse d'images multi-spectrales, au sein desquelles les valeurs – vectorielles – des pixels ne sont pas toutes comparables. Des travaux préliminaires visent à explorer les hiérarchies dans de tels espaces de valeurs, par exemple en développant des « extensions multi-valuées » de structures de données classiques comme les arbres de coupes (Naegel et Passat, 2013; Passat et Naegel, 2014). La structuration non-arborescente d'une hiérarchie peut également dériver des liens structurels entre les différentes entités constitutives de l'image. En particulier, si l'on remplace les liens d'adjacence non-orientés entre les pixels, par des liens orientés, la propriété fondamentale de non-chevauchement entre les composantes d'une image disparaît, et l'on peut alors définir non plus des hiérarchies de partitions – c'est-à-dire des arbres – mais des hiérarchies de couvertures. Des travaux préliminaires sur de tels graphes orientés acycliques ont été lancés (Tankyevych et al., 2013) afin de développer des techniques capables d'injecter des informations sémantiques au cœur même des relations structurelles dans les images.

Face à la complexité des dernières générations d'images satellitaires, l'intégration de connaissances de haut-niveau dans les approches hiérarchiques d'analyse d'images tend à devenir une nécessité. L'utilisation de telles connaissances, comme par exemple des hiérarchies de concepts, au sein des approches multi-échelles est toutefois confrontée au problème de leur intégration effective dans les algorithmes. Ces derniers étant souvent basés sur des propriétés bas-niveaux de l'image, ils ne peuvent pas nécessairement être étendus pour prendre en compte ces connaissances. L'interactivité entre les méthodes et l'expert fournit alors une piste de recherche envisageable. Cette piste mène notamment au domaine de l'*Active Learning* (Crawford et al., 2013), actuellement en pleine expansion. Les compétences de l'utilisateur y sont utilisées afin de guider des algorithmes de classification vers la création de classes d'objets d'intérêt en adéquation avec les besoins applicatifs. Ces approches pourraient notamment être étendues au traitement de données multi-sources. En particulier, l'interactivité pourrait ainsi devenir multi-temporelle ou multi-résolution. Dans ce cadre, les approches basées sur le paradigme de *Domain Adaptation* pourraient permettre de propager ces connaissances à travers le temps (Bahirat et al., 2012) ou la résolution (Kurtz et al., 2012b). Cette piste – déjà considérée dans des travaux préliminaires – n'est évidemment pas la seule envisageable. à titre d'exemple, il serait aussi possible d'intégrer des connaissances permettant de guider la construction des structures hiérarchiques représentant les images, ou bien encore d'unifier des hiérarchies d'informations représentant les images, aux hiérarchies conceptuelles modélisant les connaissances de l'utilisa-

teur vis-à-vis du contenu des données.

L'extension des approches hiérarchiques à l'analyse d'images multi-résolutions met en évidence la pertinence d'employer des mécanismes de remise en cause des résultats à une résolution donnée, par ceux obtenus à une autre résolution. Par exemple, une région d'intérêt dont les frontières sont délimitées *via* une image MR peuvent être affinées par le biais d'informations extraites à partir d'une image HR ou THR. De manière similaire, la classe thématique de cette région pourrait être remise en cause et/ou spécialisée *via* des informations thématiques provenant d'images à plus hautes résolutions. Ce constat ouvre de nouvelles pistes de recherche sur des mécanismes collaboratifs de remise en cause des résultats par segmentation et classification croisées, dont l'utilisation d'un post-processus ascendant de correction des régions d'une image extraite lors d'une phase descendante à travers la résolution, constitue un exemple représentatif. Néanmoins, des verrous scientifiques restent à lever, tels que la définition de critères permettant de contrôler un tel processus itératif, c'est-à-dire de « démêler le vrai du faux », mais aussi d'en gérer la convergence et la terminaison. Les ontologies multi-résolutions constituent une solution potentielle pour modéliser des connaissances du domaine à même de fournir de tels critères. Une réflexion plus générale reste néanmoins à mener sur le rôle des connaissances du domaine dans le processus global d'extraction d'informations. En effet, si ce type d'intégration est de plus en plus courant, peu d'études présentent des avancées aisément généralisables.

5.2. Perspectives

Avec la disponibilité croissante des images satellitaires, il devient courant de considérer des séries temporelles pour analyser le développement et l'évolution des territoires. Des images à hautes résolutions spatiales et temporelles tendent notamment à devenir largement accessibles (*e.g.*, constellation de satellites SENTINEL-2), apportant un niveau de précision très supérieur aux anciennes images. Pour profiter de cette richesse, il est toutefois nécessaire d'utiliser ces données potentiellement multi-sources dans leur intégralité, alors que la plupart des approches hiérarchiques sont limitées par leur nature monosource. Dans cet article, nous avons illustré la possibilité d'étendre ces approches au traitement de données multi-résolutions, et donc multi-sources. Parallèlement, le potentiel des approches hiérarchiques pour le traitement d'images multi-temporelles a également été mis en évidence (Bruzzone et Carlin, 2006). Une perspective prometteuse consisterait alors à combiner ces différentes approches pour permettre le traitement d'ensembles hétérogènes d'images tant multi-résolutions que multi-temporelles. Cette combinaison permettrait notamment de comprendre une scène à différentes échelles et temporalités. Néanmoins, afin de permettre une telle combinaison, il est nécessaire de définir des stratégies de mise en correspondance de hiérarchies issues de différentes images. Bien que ce pro-

blème ait déjà été étudié dans le cas de la mise en correspondance de pixels, son passage au cas des régions – et *a fortiori* des hiérarchies de régions – reste un problème ouvert. Dans l'attente de sa résolution, il a toutefois été montré (Petitjean et al., 2012) comment contourner le verrou de la mise en correspondance de régions pour l'analyse d'images multi-temporelles, *via* la notion de pixels enrichis par des caractéristiques spatio-temporelles. L'extension d'une telle approche au cas des hiérarchies de régions passe par le développement de nouvelles structures de graphes, telles que les hyperarbres, où les arêtes représentent des liens temporels ou bien des liens de composition spatiale.

Le traitement des masses de données multi-sources représente un véritable défi – tant scientifique que technologique – pour le domaine de l'analyse d'images. Les extraordinaires quantités de données acquises rendent en effet difficilement applicables, dans un contexte opérationnel, l'utilisation des différents paradigmes hiérarchiques, généralement coûteux en ressources de calcul et de mémoire. Au demeurant, le développement de nouveaux projets liés aux mosaïquage d'image THR – cartographie globale (Soille, 2006), images *GigaPixels* (Gueguen et al., 2011), etc. – contribue à amplifier ce phénomène. Les structures hiérarchiques étant bien adaptées à la division des traitements, une solution pour la gestion de ces masses de données repose alors sur la parallélisation de ces approches. Des travaux récents ont d'ores et déjà démontré qu'il était possible de paralléliser l'étape de construction de structures hiérarchiques en sous-étapes indépendantes, plus légères à gérer (Wilkinson et al., 2008; Bernabe et al., 2013). De telles stratégies autorisent le déploiement des processus hiérarchiques sur des machines à architectures parallèles : super-calculateurs ou systèmes GPU. Toutefois, l'une des faiblesses de ces stratégies est de ne pas considérer le contenu des images lors de l'étape de division de l'espace des données. Ceci peut alors conduire à des représentations erronées des objets d'intérêt représentés dans les hiérarchies. Il est ainsi nécessaire de définir de nouvelles stratégies considérant, au moment même de la parallélisation, le contenu et les frontières des objets d'intérêt de l'image. Des approches de segmentation telles que celle présentée dans cet article pourraient notamment être utilisées à de telles fins, démontrant – s'il en était encore besoin – la pertinence des approches hiérarchiques à tous les niveaux d'analyse des images.

Références

- Akcaay, H. G., Aksoy, S., 2008. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Transactions on Geoscience and Remote Sensing* 46 (7), 2097–2111.
- Baatz, M., Schape, A., 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. Dans : *Proceedings of the Angewandte Geographische Informations Symposium*. pp. 12–23.
- Bahirat, K., Bovolo, F., Bruzzone, L., Chaudhuri, S., 2012. A novel domain adaptation Bayesian classifier for updating land-cover maps with class differences in source and target do-

- mains. *IEEE Transactions on Geoscience and Remote Sensing* 50 (7), 2810–2826.
- Benediktsson, J. A., Chanussot, J., Moon, W. M., 2012. Very High-Resolution remote sensing: Challenges and opportunities. *Proceedings of the IEEE* 100 (6), 1907–1910.
- Benediktsson, J. A., Pesaresi, M., Amason, K., 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing* 41 (9), 1940–1949.
- Bernabe, S., López, S., Plaza, A., Sarmiento, R., 2013. GPU implementation of an automatic target detection and classification algorithm for hyperspectral image analysis. *IEEE Geoscience and Remote Sensing Letters* 10 (2), 221–225.
- Bruzzzone, L., Carlin, L., 2006. A multilevel context-based system for classification of Very High Spatial Resolution images. *IEEE Transactions on Geoscience and Remote Sensing* 44 (9), 2587–2600.
- Bruzzzone, L., Cossu, R., Vernazza, G., 2002. Combining parametric and non-parametric algorithms for a partially unsupervised classification of multitemporal remote-sensing images. *Image Fusion* 3 (1), 289–297.
- Collet, C., Caloz, R., Rochon, G., 2001. *Précis de télédétection : Tome 3, Traitements numériques d'images de télédétection. Vol. 3. Presses de l'Université du Québec.*
- Crawford, M. M., Tuia, D., Yang, H. L., 2013. Active learning: Any value for classification of remotely sensed data? *Proceedings of the IEEE* 101 (3), 593–608.
- de Lussy, F., Kubik, P., Greslou, D., Pascal, V., Gigord, P., Cantou, J. P., 2005. Pleiades-HR image system products and geometric accuracy. Dans : *Proceedings of the ISPRS Hannover Workshop on High-Resolution Earth Imaging for Geospatial Information. Vol. 1. pp. 50–57.*
- Dou, W., Chen, Y., Li, X., Sui, D. Z., 2007. A general framework for component substitution image fusion: An implementation using the fast image fusion method. *Computers & Geosciences* 33 (2), 219–228.
- Felzenszwalb, P., Huttenlocher, D., 2004. Efficient graph-based image segmentation. *International Journal of Computer Vision* 59 (2), 167–181.
- Forestier, G., Puissant, A., Wemmert, C., Gançarski, P., 2012. Knowledge-based region labeling for remote sensing image interpretation. *Computers, Environment and Urban Systems* 36 (5), 470–480.
- Gaetano, R., Scarpa, G., Poggi, G., 2009. Hierarchical texture-based segmentation of multiresolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 47 (7), 2129–2141.
- Garrido, L., Salembier, P., Garcia, D., 1998. Extensive operators in partition lattices for image sequence analysis. *Signal Processing* 66 (2), 157–180.
- Goetz, A. F. H., 2009. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment* 113 (1), 5–16.
- Goffe, R., Brun, L., Damiand, G., 2011. Tiled top-down combinatorial pyramids for large images representation. *International Journal of Imaging Systems and Technology* 21 (1), 28–36.
- Gruber, T. R., 1995. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies* 43 (5), 907–928.
- Gueguen, L., Pesaresi, M., Soille, P., 2011. An interactive image mining tool handling gigapixel images. Dans : *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. pp. 1581–1584.*
- Gueguen, L., Soille, P., Pesaresi, M., 2010. Differential morphological decomposition segmentation: A multi-scale object based image description. Dans : *Proceedings of the IEEE International Conference on Pattern Recognition. pp. 938–941.*
- Inglada, J., Michel, J., 2009. Qualitative spatial reasoning for High-Resolution remote sensing image analysis. *IEEE Transactions on Geoscience and Remote Sensing* 47 (2), 599–612.
- Jacobs, A., 2009. The pathologies of big data. *Communications of the ACM* 52 (8), 36–44.
- Kopanas, I., Avouris, N. M., Daskalaki, S., 2002. The role of domain knowledge in a large scale data mining project. Dans : *Proceedings of the Hellenic Conference on AI. Vol. 2308 de Lecture Notes in Computer Science. Springer, pp. 746–746.*
- Krieger, G., Gebert, N., Moreira, A., 2008. Multidimensional waveform encoding: A new digital beamforming technique for synthetic aperture radar remote sensing. *IEEE Transactions on Geoscience and Remote Sensing* 46 (1), 31–46.
- Kurtz, C., Gançarski, P., Passat, N., Puissant, A., 2013. A hierarchical semantic-based distance for nominal histogram comparison. *Data & Knowledge Engineering* 87, 206–225.
- Kurtz, C., Passat, N., Gançarski, P., Puissant, A., 2011a. Hierarchical segmentation of multiresolution remote sensing images. Dans : *Proceedings of the International Symposium on Mathematical Morphology. Vol. 6671 de Lecture Notes in Computer Science. Springer, pp. 343–354.*
- Kurtz, C., Passat, N., Gançarski, P., Puissant, A., 2011b. Multiresolution region-based clustering for urban analysis. *International Journal of Remote Sensing* 31 (22), 5941–5973.
- Kurtz, C., Passat, N., Gançarski, P., Puissant, A., 2012a. Extraction of complex patterns from multiresolution remote sensing images: A hierarchical top-down methodology. *Pattern Recognition* 45 (2), 685–706.
- Kurtz, C., Puissant, A., Passat, N., Gançarski, P., 2011c. An interactive approach for extraction of urban patterns from multisource images. Dans : *Proceedings of the IEEE Joint Urban Remote Sensing Event. pp. 321–324.*
- Kurtz, C., Puissant, A., Passat, N., Gançarski, P., 2012b. Domain adaptation for the extraction of complex urban patterns from multiresolution satellite images. Dans : *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. pp. 1773–1776.*
- Kurtz, C., Stumpf, A., Malet, J.-P., Gançarski, P., Puissant, A., Passat, N., 2014. Hierarchical extraction of landslides from multiresolution remotely sensed optical images. *ISPRS Journal of Photogrammetry and Remote Sensing* (sous presse).
- Lindeberg, T. (Ed.), 2004. *Scale-Space Theory in Computer Vision. Kluwer Academic Publishers.*
- Liu, X., 2008. Airborne LiDAR for DEM generation: Some critical issues. *Progress in Physical Geography* 32 (1), 31–49.
- Mallat, S. G., 1996. Wavelets for a vision. *Proceedings of the IEEE* 84 (4), 604–614.
- Monasse, P., Guichard, F., 2000. Scale-space from a level lines tree. *Journal of Visual Communication and Image Representation* 11 (2), 224–236.
- Mongus, D., Zalik, B., 2014. Computationally efficient method for the generation of a digital terrain model from airborne LiDAR data using connected operators. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (sous presse).
- Montanvert, A., Meer, P., Rosenfeld, A., 1991. Hierarchical image analysis using irregular tessellations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (4), 307–316.
- Naegel, B., Passat, N., 2013. Toward connected filtering based on component-graphs. Dans : *Proceedings of the International Symposium on Mathematical Morphology. Vol. 7883 de Lecture Notes in Computer Science. Springer, pp. 350–361.*
- Najman, L., Schmitt, M., 1996. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (12), 1163–1173.

- Najman, L., Talbot, H. (Eds.), 2008. Morphologie mathématique 1 : approches déterministes. Hermès-Lavoisier.
- Passat, N., Naegel, B., 2014. Component-trees and multivalued images: Structural properties. *Journal of Mathematical Imaging and Vision* (sous presse).
- Pesaresi, M., Benediktsson, J. A., 2001. A new approach for the morphological segmentation of High-Resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing* 39 (2), 309–320.
- Petitjean, F., Kurtz, C., Passat, N., Gançarski, P., 2012. Spatio-temporal reasoning for the classification of satellite image time series. *Pattern Recognition Letters* 33 (13), 1805–1815.
- Puissant, A., Weber, C., 2002. The utility of Very High Spatial Resolution images to identify urban objects. *Geocarto International* 17 (1), 33–44.
- Salembier, P., Garrido, L., 2000. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Transactions on Image Processing* 9 (4), 561–576.
- Salembier, P., Oliveras, A., Garrido, L., 1998. Antiextensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing* 7 (4), 555–570.
- Sheth, A., Henson, C., Sahoo, S. S., 2008. Semantic Sensor Web. *IEEE Internet Computing* 12 (4), 78–83.
- Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (8), 888–905.
- Soille, P., 2006. Morphological image compositing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (5), 673–683.
- Soille, P., Pesaresi, M., 2002. Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Transactions on Geoscience and Remote Sensing* 40 (9), 2042–2055.
- Sun, W., Heidt, V., Gong, P., Xu, G., 2003. Information fusion for rural land-use classification with High-Resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing* 41 (4), 883–890.
- Tankyevych, O., Talbot, H., Passat, N., 2013. Semi-connections and hierarchies. Dans : *Proceedings of the International Symposium on Mathematical Morphology*. Vol. 7883 de *Lecture Notes in Computer Science*. Springer, pp. 157–168.
- Vilaplana, V., Marques, F., Salembier, P., 2008. Binary partition trees for object detection. *IEEE Transactions on Image Processing* 17 (11), 2201–2216.
- Vincent, L., Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (6), 583–598.
- Wang, F., 1991. Integrating GIS's and remote sensing image analysis systems by unifying knowledge representation schemes. *IEEE Transactions on Geoscience and Remote Sensing* 29 (4), 656–664.
- Wemmert, C., Puissant, A., Forestier, G., Gançarski, P., 2009. Multiresolution remote sensing image clustering. *IEEE Geoscience and Remote Sensing Letters* 6 (3), 533–537.
- Wilkinson, M. H. F., Gao, H., Hesselink, W. H., Jonker, J. E., Meijster, A., 2008. Concurrent computation of attribute filters on shared memory parallel machines. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (10), 1800–1813.
- Witharana, C., Civco, D. L., Meyer, T. H., 2013. Evaluation of pansharpening algorithms in support of Earth observation based rapid-mapping workflows. *Applied Geography* 37 (1), 63–87.