

Apprentissage de règles floues pour caractériser des objets d'intérêt dans une image de télédétection

Bruno Belarte*¹, Cédric Wemmert*
Germain Forestier**, Christiane Weber***
Manuel Grizonnet****

*Université de Strasbourg, ICube, UMR CNRS 7357, Strasbourg, France
{belarte,wemmert}@unistra.fr

**Université de Haute Alsace, MIPS EA 2332, Mulhouse, France
germain.forestier@uha.fr

***Université de Strasbourg, LIVE, UMR CNRS 7362, Strasbourg, France
christiane.weber@live-cnrs.unistra.fr

****CNES, DCT/SI/AP

18, avenue Edouard Belin, 31401 Toulouse Cedex 09
manuel.grizonnet@cnes.fr

Résumé. Les nouveaux capteurs satellitaires permettent l'acquisition d'images d'un très haut niveau de détail à des cadences élevées, produisant ainsi une importante masse de données. Le traitement manuel de ces données étant devenu impossible, de nouveaux outils sont nécessaires afin de les traiter automatiquement. Dans ce cadre, des algorithmes de segmentation efficaces sont nécessaires pour extraire des objets d'intérêt de ces images. Cependant, les segments produits ne correspondent généralement pas aux objets d'intérêt. Dans cet article, nous proposons de changer le niveau d'abstraction afin d'interpréter les objets d'intérêt comme des objets composés par des segments. Pour cela, nous avons mis en place un processus d'apprentissage multi-niveaux, basé sur des connaissances expertes, dans le but d'apprendre des règles de compositions définissant des objets d'intérêt. Pour gérer l'imprécision relative à l'analyse d'images de télédétection nous proposons d'utiliser la logique floue afin de modéliser les règles de composition. La méthode proposée est validée sur des données de synthèse ainsi que sur des données réelles.

1 Introduction

L'évolution croissante de la résolution d'acquisition des images dans tous les domaines (imagerie médicale, télédétection, etc.) permet d'obtenir des images d'un très haut niveau de détail. Pour pouvoir analyser ce type d'images, des méthodes dites *orientées régions* ont été développées. Ce type d'approches offre la possibilité de s'intéresser aux propriétés (radiométriques, structurelles, spatiales, etc.) de ces régions.

1. Les travaux présentés ont été financés par le Centre National des Études Spatiales et la Région Alsace.

Apprentissage de règles floues

L'utilisation de connaissance experte est fréquente en extraction d'informations. Dans la classification non supervisée par contraintes, la connaissance est exprimée sous la forme de deux relations signifiant que deux objets *doivent* ou *ne peuvent pas* appartenir à la même classe. Wagstaff et al. (2001) proposent une version contrainte de l'algorithme K-MEANS utilisant de telles contraintes pour biaiser l'affectation des objets aux classes. À chaque étape, l'algorithme essaie de s'accorder avec les contraintes données. Bilenko et al. (2004) utilisent ces contraintes pour apprendre une fonction de distance biaisée par les connaissances. Kumar et Kummamuru (2008) introduisent un autre type de connaissances, les comparaisons relatives, à travers un algorithme de classification non supervisée. Leur étude expérimentale démontre que l'algorithme proposé est plus précis et robuste avec ces relations qu'avec les contraintes présentées plus haut. Pedrycz (2004) introduit plusieurs sortes de connaissances du domaine telles que la supervision partielle, les conseils basés sur la proximité ou sur l'incertitude. L'auteur propose plusieurs solutions pour exploiter et intégrer ces connaissances dans un algorithme Fuzzy C-MEANS.

Afin de modéliser ces connaissances implicites, l'utilisation d'ontologie devient de plus en plus courante. Une ontologie (Gruber, 1995) est une spécification abstraite, une vue simplifiée du monde représentée dans un but précis. Une ontologie définit un ensemble de concepts, leurs caractéristiques et les relations entre ces concepts. L'utilisation d'ontologies est courante dans les systèmes d'informations géographiques (Fonseca et al., 2002). De plus, un intérêt croissant a été porté aux ontologies spécialisées pour l'analyse d'image (Bittner et Winter, 1999) depuis ces dix dernières années. La plupart des méthodes (Forestier et al., 2008, 2012; Maillot et Thonnat, 2008; Mezaris et al., 2004; Durand et al., 2007) formalisent les concepts pouvant être présents dans une image puis proposent une analyse sémantique de celle-ci en cherchant à identifier des représentants de ces concepts dans l'image.

À partir d'une description faite par l'expert des objets d'intérêt, il est possible de caractériser les régions extraites. Cependant, le niveau de détail des images est désormais tel que les algorithmes de segmentation usuels ne sont pas capables de produire des régions correspondant à des objets d'intérêt pour l'expert, rendant caduque l'utilisation de connaissances expertes. En effet, les objets d'intérêt se retrouvent découpés en plusieurs segments. C'est dans ce contexte que s'inscrivent nos travaux. Nous proposons dans cet article un algorithme d'apprentissage de règles de composition qui vise à apprendre comment sont composés les objets d'intérêt dans une image en faisant l'hypothèse qu'elle est sur-segmentée. Ces travaux poursuivent les travaux de Belarte et Wemmert (2012). L'originalité de cette approche repose sur le fait de faire le lien entre la connaissance de l'expert et une image segmentée, non pas en considérant un segment comme un objet d'intérêt mais en considérant un objet d'intérêt comme composé par un ou plusieurs segments. Le reste de l'article est organisé de la manière suivante : la section 2 présente une méthode d'analyse d'image guidée par la connaissance. La section 3 introduit les notions relatives à la logique floue et à l'apprentissage de règles. La méthode d'apprentissage de règles de composition proposée est présentée dans la section 4. La section 5 présente les principaux résultats obtenus. Enfin, la section 6 conclut et évoque des perspectives de travaux futurs.

2 Analyse d'image guidée par la connaissance

La méthode d'analyse d'images que nous proposons repose sur l'identification d'objets d'intérêt à partir de leur définition par un expert du domaine. Pour cela, il est nécessaire de formaliser la représentation des objets d'intérêt et de la base de connaissances. Dans un premier temps, nous allons présenter notre formalisation des objets d'intérêt avant de décrire la notion de concept et l'organisation hiérarchique de ces concepts.

2.1 Objet d'intérêt

Un objet d'intérêt est un objet ayant une réalité sémantique dans le domaine d'expertise. Ils peuvent être de deux types. Les *objets d'intérêt simples* représentent la granularité la plus fine pour l'expert. Les *objets d'intérêt composés* sont des agrégats d'objets simples et/ou d'autres objets composés. Dans le contexte de ces travaux, un objet simple peut être par exemple une maison, un arbre, une route ou encore une étendue d'eau ; un objet composé peut être par exemple un îlot pavillonnaire, un parc, une zone d'activité ou encore un aéroport. Un objet d'intérêt est représenté par sa forme géométrique (le polygone englobant cet objet) et par trois types de propriétés numériques définies comme suit :

Attributs : Un attribut propre est une propriété qui ne dépend que de l'objet en lui-même. Ces propriétés décrivent notamment la signature radiométrique de l'objet (valeur moyenne des bandes, indices NDVI, NDWI2, SBI, CI) et sa forme (taille, élongation). Les indices radiométriques sont définis par les équations 1, 2, 3 et 4, où *rouge*, *bleu*, *vert* et *pir* (proche infrarouge) représentent les valeurs moyennes de chacune des quatre bandes spectrales de l'image.

$$NDVI = \frac{pir - rouge}{pir + rouge} \quad (1)$$

$$NDWI2 = \frac{vert - pir}{pir + vert} \quad (2)$$

$$CI = \frac{rouge - vert}{rouge + vert} \quad (3)$$

$$SBI = 0.2623 \cdot bleu + 0.6432 \cdot vert + 0.6302 \cdot rouge + 0.3471 \cdot pir \quad (4)$$

Relations spatiales : Une relation spatiale est une propriété qui lie l'objet courant à son voisinage. La relation d'adjacence à un objet de type A est définie par le ratio des côtés du polygone partagés avec un objet de type A sur le périmètre de l'objet courant. La relation de distance à un objet de type B est définie par la distance en mètres à l'objet de type B le plus proche de l'objet courant.

Relations de composition : Une relation de composition est une propriété qui lie l'objet courant à des objets de niveaux d'abstraction différents. La relation de composition d'un objet composé par des objets de type C est définie par le ratio de la surface de chacun des objets de type C composant l'objet courant sur la taille de l'objet courant.

2.2 Concept

À partir de cette définition d'un objet d'intérêt, nous pouvons définir la notion de concept. Un *concept* est une représentation abstraite d'un objet d'intérêt. L'expert dispose d'un grand nombre de variables linguistiques dans son domaine d'expertise. Par exemple, la propriété *élongation* est définie par les variables linguistiques *compact*, *moyennement allongé* et *très allongé*. L'expert peut formuler des règles à partir des différentes variables linguistiques. Chaque règle décrit un concept. Par exemple, le concept *arbre* est défini par la règle *petit et compact et NDVI élevé*. Dans un paysage urbain, l'expert peut définir des concepts de plus haut niveau d'abstraction. Un objet correspondant à un tel concept composé est un agrégat d'objets élémentaires ou d'autres objets composés. Un *concept composé* est défini par une règle portant sur tout type de propriétés (attributs, relations spatiales, relations de composition). Par exemple, le concept composé *îlot pavillonnaire* est défini par la règle *grand et composé de pavillons et composé de végétation et entouré par de la route*.

La manière la plus simple de représenter un concept est de le représenter sous la forme d'une liste de contraintes sur les propriétés d'un objet. De là, émergent plusieurs questions. Comment formaliser ces contraintes pour qu'elles puissent prendre en compte l'imprécision inhérente au traitement d'images ? Quel formalisme utiliser pour que l'expert puisse produire des règles concises et réutilisables ? Nous avons choisi de représenter nos contraintes par des ensembles flous car ils permettent de gérer l'imprécision tout en étant faciles à mettre en place. Ceux-ci sont présentés dans la section 3.

2.3 Base de connaissances

L'expert peut définir un concept qui est un cas particulier d'un autre concept. Par exemple, le concept *étang* est un cas particulier du concept *étendue d'eau*, ou, le concept *végétation* est une généralisation du concept *pelouse*. Pour pouvoir ordonner les concepts en conséquence, ils sont placés dans une hiérarchie dans laquelle une feuille est un concept particulier et chaque noeud est une généralisation de ses sous-arbres. Un exemple de hiérarchie est donné dans la figure 1. Dans cet exemple, nous pouvons voir des concepts simples et un concept composé ainsi que les relations de composition définissant le concept *îlot résidentiel*. L'extraction des objets composés repose sur un algorithme de croissance de régions (Adams et Bischof, 1994) adapté aux graphes d'objets d'intérêt.

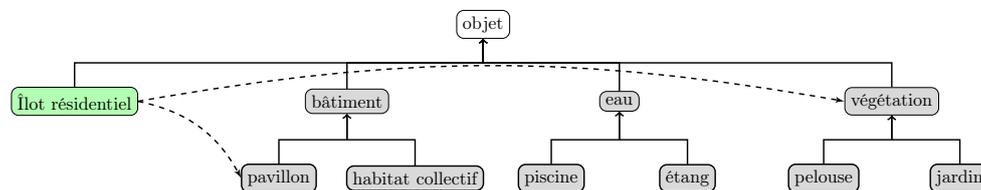


FIG. 1 – Exemple d'une hiérarchie de concepts. En gris, les concepts simples, en vert un concept composé. Les pointillés représentent les relations de composition.

2.4 Limites de l'analyse d'images THRS

Dans la pratique, il n'est pas possible d'utiliser une telle approche pour analyser des images de télédétection. Des tests ont permis de mettre en avant le problème suivant : les objets analysés (produits par la segmentation) ne correspondent pas à des objets sémantiques (décrits par les concepts). Ceci est dû à un phénomène de sur-segmentation ou de sous-segmentation ; la classification des objets d'intérêt est très fortement biaisée par cette imprécision. Un exemple est donné dans la figure 2. Nous cherchons à identifier les pavillons. Le pavillon entouré en rouge sur la vérité terrain n'est tout simplement pas identifié. Il s'avère que l'algorithme de segmentation l'a décomposé en plusieurs segments ce qui le rend non identifiable par la base de connaissances (contrainte de taille non satisfaite). Le pavillon à droite est correctement identifié mais le pavillon à gauche n'est pas détecté.

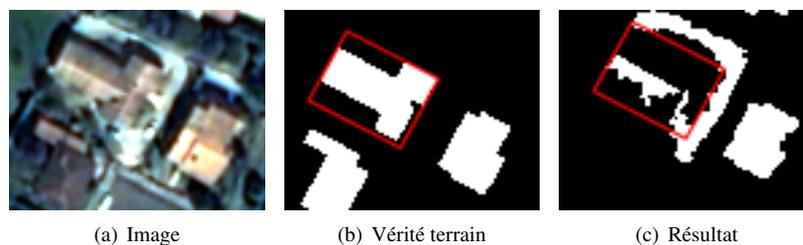


FIG. 2 – *Illustration des résultats obtenus par un algorithme de classification basé sur la connaissance. Nous pouvons constater que la maison de droite est correctement reconnue, que la maison en bas à gauche n'est pas reconnue et la maison entourée de rouge n'est pas reconnue intégralement.*

De ce fait, l'extraction des objets composés n'est pas réalisable puisqu'elle repose sur un étiquetage correct du graphe. En effet, l'utilisation de connaissances décrivant des objets sémantiques repose sur la bonne détection de ces objets dans l'image. Or les algorithmes de segmentation classiques ne sont pas capables de retrouver de tels objets. Un algorithme de segmentation repose sur l'identification de zones homogènes au niveau image et non sur l'identification d'objets sémantiques. Dans une image à très haute résolution spatiale (de type QUICKBIRD ou PLEIADE) les objets d'intérêt pour l'expert se retrouvent découpés en plusieurs segments. Il n'y a pas de relation bijective entre les objets géographiques et les segments. La connaissance experte n'est de fait pas utilisable en l'état pour l'analyse d'image de télédétection à très haute résolution spatiale.

La solution que nous proposons pour remédier à cette limitation est de baisser le niveau d'abstraction. En considérant les objets sémantiques non plus comme des objets élémentaires mais comme des objets composés, notre approche reste valide. Il suffit de décrire un objet d'intérêt par une règle de composition. Cependant, l'expert n'est pas capable de donner une telle règle de composition car cette règle va dépendre du type et de la résolution de l'image, de l'algorithme de segmentation utilisé, de l'orientation des objets ou du moment d'acquisition de l'image. Avec une telle combinaison de facteurs, il n'est pas possible pour l'expert de donner une règle générique décrivant un concept. Nous proposons d'utiliser un algorithme d'apprentissage pour apprendre une règle de composition à partir d'exemples donnés par l'expert.

3 La logique floue comme moteur de raisonnement

La théorie des ensembles flous a été introduite par Zadeh (1965). C'est une théorie mathématique qui étend la théorie ensembliste classique en permettant la gestion de l'imprécision et de l'erreur. Novák (2012) propose un état de l'art des différentes théories et de leurs évolutions. La logique floue (Zadeh, 1975) permet de modéliser la connaissance à l'aide de variables linguistiques. Elle est notamment utilisée depuis la fin des années 70 (Mamdani et Assilian, 1975; Mamdani, 1977) comme moteur de raisonnement dans des systèmes experts. Lee (1990a,b) propose une revue de différentes applications de contrôleurs flous ainsi que des différentes stratégies et mécanismes de raisonnement.

La logique floue peut être utilisée par des algorithmes de classification. Lorsqu'il n'existe pas de connaissance *a priori*, il est nécessaire d'apprendre des règles à partir d'exemples. Plusieurs familles de méthodes existent. Parmi celles-ci nous pouvons notamment citer les méthodes basées sur les réseaux de neurones et celles basées sur les algorithmes évolutionnaires. Dans les méthodes basées sur les réseaux de neurones (Chakraborty et Pal, 2004; Cpalka, 2009; Wu et al., 2010; Suresh et Subramanian, 2011; Sebari et He, 2013) les règles sont modélisées par des neurones et les fonctions d'appartenances par des neurones ou par des connexions. Les neurones de la couche d'entrée représentent les variables et les neurones de la couche de sortie réalisent le processus de défloutage. Les méthodes basées sur des algorithmes évolutionnaires (Castillo et al., 2001; Ishibuchi et Yamamoto, 2004; Fernández et al., 2010) considèrent un ensemble de règles comme population et font évoluer cette population pour trouver un ensemble de règles minimal. Il existe des méthodes hybrides (Russo, 2000) utilisant à la fois des réseaux de neurones et des algorithmes évolutionnaires. Les méthodes précédentes modifient la forme des fonctions d'appartenances ou la composition des règles au cours des processus d'apprentissage. D'autres méthodes (Nozaki et al., 1996; Ishibuchi et Nakashima, 2001; Nakashima et al., 2007) font varier le poids associé aux règles ou aux fonctions d'appartenances plutôt que de modifier directement la forme ou la composition des règles.

La logique floue peut être utilisée dans l'analyse d'images de télédétection. Ke et al. (2008) utilisent un classifieur flou sur une segmentation multi échelle pour extraire et surveiller la végétation urbaine. Gueguen et al. (2012) utilisent la logique floue et le filtrage morphologique pour caractériser les bâtiments sans toit. Laha et al. (2006) utilisent une carte autoadaptative de Kohonen pour générer des prototypes qui sont traduits en règles floues dans un processus de classification basé pixel. Sebari et He (2013) utilisent des méthodes d'analyses d'images orientées objets. Pour chaque propriété sélectionnée deux fonctions d'appartenance sont apprises et combinées en règles dans le but de classer les objets segmentés.

Dans la suite de cette section nous rappelons les bases de la théorie des ensembles flous et de la logique floue.

3.1 Théorie des ensembles flous

Dans la théorie ensembliste classique, le degré d'appartenance d'un élément à un ensemble est donné par une valeur binaire : l'élément appartient ou non à l'ensemble. Le degré d'appartenance à un ensemble flou est décrit à l'aide d'une fonction d'appartenance retournant une valeur réelle dans l'intervalle $[0, 1]$, 1 signifiant que l'élément appartient totalement à l'ensemble, 0 signifiant que l'élément n'appartient pas du tout à l'ensemble. Toutes les valeurs intermédiaires permettent de nuancer le degré d'appartenance.

Une fonction d'appartenance est définie par deux ensembles, le noyau et le support. Le *noyau* est l'ensemble des éléments appartenant totalement à l'ensemble (dont le degré de vérité est égal à 1). Le *support* est l'ensemble des éléments appartenant au moins un peu à l'ensemble (dont le degré de vérité est strictement positif). Par conséquent, un élément peut appartenir à deux ensembles flous en même temps mais à des degrés différents (par exemple, un élément peut être *petit* avec un degré de 0.86 et *moyen* avec un degré de 0.27).

Définition formelle. Soit E l'espace des données, $A \subset E$ un ensemble flou, n_A le noyau et s_A le support de l'ensemble A , x un élément de E , μ_A la fonction d'appartenance modélisant l'ensemble flou A . Le noyau et le support sont définis dans les équations 5 et 6.

$$n_A = \{x \in E | \mu_A(x) = 1\} \quad (5)$$

$$s_A = \{x \in E | \mu_A(x) > 0\} \quad (6)$$

Opérations. Les opérateurs classiques d'union, d'intersection et de complément sont étendus aux ensembles flous. L'intersection est modélisée par une T-norme et l'union par une T-conorme. Dans son papier original, (Zadeh, 1965) propose comme T-norme l'opérateur minimum et comme T-conorme l'opérateur maximum. Soit E le domaine, $A \subset E$ et $B \subset E$ deux ensembles flous. Les équations 7, 8 et 9 illustrent ces opérations.

$$\mu_{A \cap B} = \min(\mu_A, \mu_B) \quad (7)$$

$$\mu_{A \cup B} = \max(\mu_A, \mu_B) \quad (8)$$

$$\mu_{\bar{A}} = 1 - \mu_A \quad (9)$$

Comme pour les opérations ensemblistes classiques, les opérations d'union et d'intersection sont commutatives et associatives, l'intersection est distributive par rapport à l'union. Cependant, le principe du tiers exclu est contredit ($\mu_A \cup \mu_{\bar{A}} \neq 1$). D'autre part, un élément peut appartenir à un ensemble et à son complément (mais à des degrés différents).

Un ensemble flou étant étroitement lié à la fonction d'appartenance le représentant, nous utiliserons par la suite indifféremment les termes *ensemble flou* et *fonction d'appartenance*.

3.2 Logique floue

La logique floue (Zadeh, 1975) est une technique de raisonnement en intelligence artificielle. Elle repose sur la théorie des ensembles flous présentée dans la section précédente. Le principe de la logique floue est basé sur l'utilisation de variables linguistiques pour modéliser la connaissance. Plutôt que de définir une contrainte par un intervalle de valeurs, elle est définie par une fonction d'appartenance. Ainsi, une variable telle que la *taille* va pouvoir prendre comme valeur *petit* ou *grand*. L'expert peut alors associer des ensembles flous à des variables linguistiques du domaine qu'il sera possible d'utiliser dans un second temps pour former des règles. Le fait que ces variables linguistiques soient basées sur des ensembles flous permet de gérer l'imprécision des mesures à travers l'utilisation du degré d'appartenance.

Règles. Les variables linguistiques peuvent être combinées à l'aide d'opérateurs logiques **et**, **ou** et **non**. L'opérateur **et** est modélisé par une T-norme, l'opérateur **ou** est modélisé par

une T-conorme, l'opérateur **non** est modélisé par le complément, ceci de manière similaire aux opérateurs de la théorie des ensembles flous. Parmi les T-normes les plus utilisées dans la littérature, nous pouvons citer l'intersection ensembliste de Zadeh (équation 7). Parmi les T-conormes les plus utilisées dans la littérature, nous pouvons citer l'union ensembliste de Zadeh (équation 8). L'opérateur **non** est toujours modélisé par le complément (équation 9). Il existe d'autres opérateurs dans la littérature (Lee, 1990a,b) tels que l'union et l'intersection probabiliste. En combinant différentes variables linguistiques, l'expert peut former des règles cohérentes et complexes. De par leurs définitions linguistiques, ces règles sont compréhensibles par un être humain. Leurs traductions par des fonctions d'appartenance les rendent compréhensibles par un algorithme mathématique.

Soit *allongé*, *brillant*, *petit* et *moyen* quatre variables linguistiques. Un exemple de règle combinant ces fonctions à l'aide des différents opérateurs est donné par la formule suivante :

$$r = \textit{allongé et non brillant et (petit ou moyen)}$$

La transformation de cette règle en expression mathématique utilisant les opérateurs ensemblistes de Zadeh est donnée dans l'équation 10.

$$\mu_r = \min(\mu_{\textit{allongé}}, \min(1 - \mu_{\textit{brillant}}, \max(\mu_{\textit{petit}}, \mu_{\textit{moyen}}))) \quad (10)$$

À chaque règle est associée une valeur de sortie. Cette valeur peut être une variable numérique ou un ensemble flou. Il est alors possible d'exprimer une base de règles sous la forme d'une liste de règles d'inférence, où A, B et N sont des variables linguistiques du même domaine :

- Si règle 1 alors A
- Si règle 2 alors B
- ...
- Si règle n alors N

Évaluation. À partir de cette base de règles, il est maintenant possible d'évaluer le système à l'aide d'un moteur de raisonnement basé sur la logique floue. Les variables du système sont transformées en variables linguistiques (processus de fuzzyfication, ou floutage) pour pouvoir être évaluées par le moteur de raisonnement. Chaque règle est ensuite évaluée. Les résultats de chaque évaluation sont agrégés à l'aide de l'opérateur maximum. La dernière étape du processus consiste à attribuer au résultat la bonne variable linguistique (processus de defuzzyfication, ou défloutage). Pour cela, il existe un grand nombre de méthodes dans la littérature (Lee, 1990a,b). Les deux plus utilisées sont le centre de gravité (*CdG*, équation 11) et la moyenne des maximums (*MM*, équation 12, où $S = \{y_x \in E, \mu(y_x) = \max_{y \in E}(\mu(y))\}$, E étant le domaine de définition de la variable de sortie). Soit d la variable de décision retournée par l'algorithme de raisonnement.

$$d_{CdG} = \frac{\int y \cdot \mu(u) \cdot dy}{\int \mu(u) \cdot dy} \quad (11)$$

$$d_{MM} = \frac{\int_S y \cdot dy}{\int_S dy} \quad (12)$$

Le processus de défloutage va donner la valeur de sortie de l'algorithme de raisonnement à partir de l'agrégat d'ensembles flous obtenu lors de l'évaluation. Cette valeur peut maintenant être utilisée dans un processus de décision.

3.3 Application à la télédétection

À partir des notions dans cette section, nous disposons de tous les outils nécessaires pour analyser des images de télédétection à l'aide de la connaissance de l'expert du domaine. Nous allons adapter ces outils pour traiter les images et les mettre en pratique.

Modélisation des données. La donnée brute que nous voulons analyser est une image de paysage urbain. Pour pouvoir l'exploiter avec un moteur de raisonnement basé sur la logique floue, nous devons transformer cette donnée brute en donnée utilisable par l'algorithme. Nous nous intéressons à une analyse orientée objet de l'image. À l'aide d'un algorithme de segmentation, l'image est transformée en segments. Un segment correspond à un objet élémentaire.

Chaque objet est défini par un polygone (caractérisant son contour) et par un ensemble de propriétés (radiométriques et structurelles) calculées à partir de l'image. Ces objets élémentaires sont ensuite placés dans un graphe d'adjacence. Un noeud correspond à un objet, un arc correspond à une relation d'adjacence entre deux objets. Deux objets sont adjacents s'ils partagent une arête commune dans l'image. Ainsi les données contenues dans le graphe correspondent à la définition des objets d'intérêt proposée en section 2.1.

L'objectif de l'analyse d'image est d'identifier les objets d'intérêt correspondant aux concepts définis par l'expert. Pour cela, nous allons adapter le moteur de raisonnement flou pour qu'il génère en sortie l'étiquette du concept à attribuer à l'objet évalué.

Moteur d'inférence flou pour l'analyse d'objets d'intérêt. La donnée de base est un objet d'intérêt. Les ensembles flous et les règles sont renseignés par l'expert du domaine et placés dans une base de connaissances. Cette base de connaissances contient la hiérarchie de concepts. Le moteur de raisonnement flou va servir à raffiner le concept attribué à chaque objet d'intérêt. Toutes les règles ne sont pas évaluées en même temps. Seules les règles correspondant aux concepts raffinant directement le concept attribué à l'objet courant (descendants directs dans la hiérarchie) sont évaluées. Une fois la nouvelle étiquette trouvée et attribuée à l'objet courant, le processus itère tant que le concept attribué a des descendants dans la hiérarchie ou que la variable de sortie ne permet pas d'attribuer une nouvelle étiquette (*i.e.* la valeur de sortie est inférieure à un seuil et ne permet pas de prendre une décision).

Les opérateurs flous utilisés sont les opérateurs ensemblistes (*i.e.* minimum et maximum). L'opérateur de défloutage a du être adapté car dans notre cas la variable de sortie n'est pas ordonnée. Nous avons choisi de prendre une variante de l'opérateur moyenne des maximums. Plutôt que de calculer la moyenne des maximums, nous gardons simplement le concept correspondant au score maximum. La formule utilisée est donnée par l'équation 13, où o est l'objet à étiqueter, x_j est le concept correspondant à la règle μ_j et n est le nombre de concepts évalués.

$$d_{\max} = \{x_j | \mu_j = \max(\mu_i(o)), i \in [1, n]\} \quad (13)$$

4 Apprentissage de règles de composition

Nous souhaitons apprendre comment sont composés une classe d'objets d'intérêt en fonction de la segmentation. L'expert donne des exemples de cette classe, les segments sous-jacents sont extraits, des règles sont apprises puis utilisées pour étiqueter les segments.

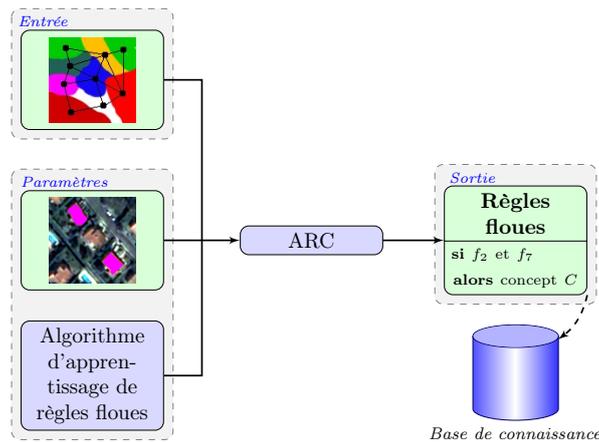


FIG. 3 – Organigramme général de la méthode. L'entrée, la sortie et les paramètres sont représentés. La base de connaissance est enrichie avec les concepts générés.

4.1 Méthode proposée : Apprentissage de règles de composition

La méthode que nous proposons est l'algorithme d'Apprentissage de Règles de Composition (abrégé ARC). ARC prend en entrée une image segmentée. Deux paramètres sont nécessaires : les exemples (fournis par l'expert) ainsi qu'un algorithme d'apprentissage de règles floues. Le déroulement général de l'algorithme est présenté dans la figure 3.

L'expert donne les exemples sous la forme d'une liste de polygones, chaque exemple étant défini par un polygone. Les objets correspondant à ces exemples doivent être extraits. Le graphe est parcouru noeud par noeud et les objets correspondant à chacun des exemples sont conservés. L'apprentissage d'une règle de composition se fait en trois étapes. Dans un premier temps, les règles portant sur les attributs radiométriques et structurels sont apprises. Chacune des règles apprises est associée à un nouveau concept qui enrichit la hiérarchie. Ces concepts sont alors utilisés pour étiqueter les segments. Une fois étiquetés, il est possible d'apprendre les règles portant sur les relations spatiales entre les segments. En effet, au début du processus les segments ne sont pas étiquetés, il n'est donc pas possible d'apprendre les règles spatiales dans la première étape. Chacune de ces nouvelles règles permet de créer un ou plusieurs concepts dans la hiérarchie. En effet, ces nouvelles règles permettent de différencier des objets étiquetés précédemment, les nouveaux concepts générés sont donc ajoutés plus bas dans la hiérarchie de concepts. Ceci est illustré dans la figure 4. Ces nouveaux concepts vont permettre de réétiqueter les segments.

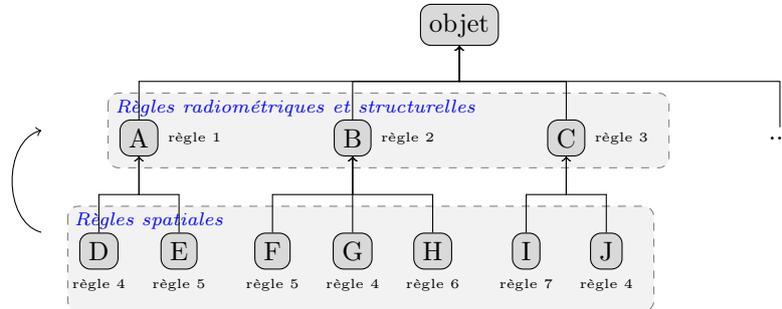


FIG. 4 – La nouvelle hiérarchie de concepts après apprentissage. Le premier niveau correspond à la sortie de la phase d'apprentissage sur les attributs propres, le second correspond à la sortie de la phase d'apprentissage sur les relations spatiales.

La troisième étape consiste en l'apprentissage de la règle de composition. Pour cela, les segments composant chaque exemple fourni par l'expert sont combinés en un objet composé. Les fonctions de compositions peuvent alors être apprises. De plus, des fonctions sur les propriétés structurelles (taille, forme, ...) sont apprises. Les propriétés radiométriques sont volontairement laissées de côté mais pourraient être intégrées par la suite, notamment pour évaluer l'homogénéité des objets composés. À partir des fonctions générées, une règle de composition unique est apprise. La figure 5 illustre le processus d'apprentissage dans son intégralité.

L'algorithme d'apprentissage de règles de composition est découpé en trois étapes d'apprentissage de règles floues simples (en pointillés dans la figure 5). La méthode d'apprentissage de règles simples n'intervient pas dans la structure de l'algorithme et peut donc être n'importe quelle méthode d'apprentissage de règles. Pour ces expériences, nous avons utilisée la méthode ARF_{CAH}^{sup} présenté dans la section suivante comme algorithme d'apprentissage de règles. Une fois l'apprentissage terminé, la nouvelle règle de composition est utilisée pour extraire les objets composés.

4.2 Apprentissage de règles floues

Nous souhaitons apprendre des règles floues de manière automatique. Nous avons montré dans l'introduction de la section 3 que la littérature est riche dans ce domaine. Nous proposons ici une méthode qui apprend indépendamment des fonctions d'appartenance pour chaque propriété et qui dans un second temps combine ces fonctions pour apprendre des règles minimalistes. La méthode que nous proposons utilise intensément l'algorithme de classification ascendante hiérarchique.

Classification Ascendante Hiérarchique (CAH). CAH est un algorithme de classification automatique (Lance et Williams, 1967). Cet algorithme repose sur une mesure de dissimilarité D entre classes qui est utilisée pour déterminer quelles sont les classes à fusionner. La dissimilarité que nous avons choisi est la distance euclidienne au carré, notée d_2^2 . Au départ, les objets

Apprentissage de règles floues

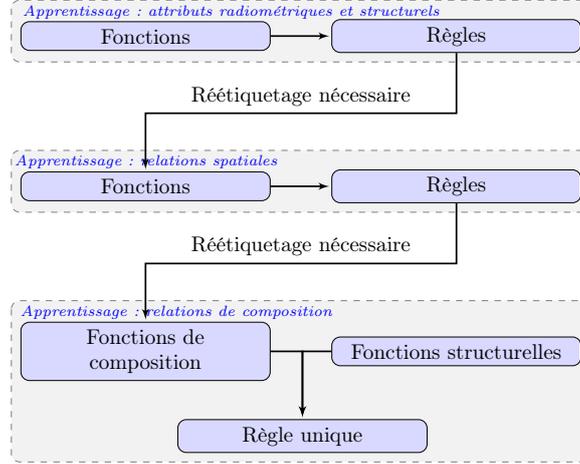


FIG. 5 – Organigramme résumant le processus d'apprentissage.

sont chacun répartis dans une classe. À chaque étape, les deux classes les plus proches sont fusionnées. La sortie de cet algorithme est un arbre de partition binaire. Le nombre de classes à conserver dépend du critère d'arrêt.

Le critère d'arrêt peut être un nombre de classes prédéfini ou un critère automatique. Dans notre cas, nous ne connaissons pas le nombre de classes recherchées, nous choisissons donc une méthode automatique. La mesure de dissimilarité de Ward (Ward, 1963) calcule la différence d'inertie intra-classe (ou inter-classe) avant et après fusion de deux classes. L'équation 14 donne la formule de la dissimilarité de Ward, où d_W est la mesure calculée, C_u et C_v sont deux classes, $n_i = \text{card}(C_i)$ et $D(C_u, C_v)$ est la dissimilarité entre deux classes. La dissimilarité utilisée est la même que celle de l'algorithme de classification hiérarchie ascendante, soit la distance euclidienne au carré d_2^2 . Lorsque la dissimilarité de Ward entre deux classes est trop élevée (*i.e.* l'augmentation de l'inertie inter-classe ou la perte d'inertie intra-classe trop élevée), la fusion de ces deux classes produirait une nouvelle classe trop hétérogène et donc non cohérente. Ce seuil critique dépend des données.

$$d_W(C_u, C_v) = \frac{n_u \cdot n_v}{n_u + n_v} \cdot D(C_u, C_v) \quad (14)$$

Apprentissage de règles floues basé sur CAH. La méthode d'apprentissage de règles que nous proposons ici est une méthode élémentaire qui repose sur l'algorithme CAH et qui ne nécessite pas d'autres paramètres que le critère de Ward. Nous appelons cette méthode ARF_{CAH} . L'objectif est d'extraire rapidement un ensemble de règles. Dans un premier temps, les fonctions d'appartenances sont apprises indépendamment pour chaque propriété. Pour cela, une CAH est effectuée pour chacune des propriétés du système. Chaque CAH produit une ou plusieurs fonctions d'appartenance. Dans un second temps, les fonctions apprises sont utilisées

pour définir les objets, chaque objet est défini par les valeurs de retour de chacune des fonctions d'appartenance apprises. Une nouvelle CAH est effectuée sur les objets nouvellement définis. Le résultat permet de générer une règle par classe.

La méthode proposée est une méthode de classification non supervisée capable d'apprendre une ou plusieurs règles à partir d'un ensemble d'objets. Cependant, elle peut aussi être utilisée comme classification supervisée. Nous pouvons distinguer les deux versions de l'algorithme. Dans le cas non supervisé, l'algorithme prend en entrée un ensemble d'objets et retourne un ensemble de règles (une par classe apprise par la CAH) séparant au mieux les différentes classes identifiées automatiquement. Nous appelons cet algorithme ARF_{CAH}^{auto} . Dans le cas supervisé, l'algorithme prend en entrée un ensemble d'exemples et retourne une règle unique décrivant au mieux l'ensemble des exemples. Les fonctions d'appartenances sont toujours apprises automatiquement, mais comme nous ne voulons qu'une seule règle, il n'y a pas de critère d'arrêt lors de l'apprentissage des règles (la CAH continue jusqu'à ce qu'il ne reste qu'une seule classe). Nous appelons cet algorithme ARF_{CAH}^{sup} .

Évaluation. Pour tester notre méthode, nous choisissons d'utiliser les jeux de données de l'UCI². Les jeux de données testés sont résumés dans le tableau 1. Nous testons notre méthode en tant que classification supervisée et classification non supervisée. Notre méthode étant basée sur l'algorithme CAH, nous évaluons aussi cet algorithme pour nous y comparer. Nous évaluons la CAH avec différentes valeurs du critère de Ward, ARF_{CAH}^{auto} avec différentes valeurs du critère de Ward pour l'apprentissage des fonctions et des règles, ARF_{CAH}^{sup} avec différentes valeurs du critère de Ward pour l'apprentissage des fonctions. Nous choisissons de comparer ARF_{CAH}^{sup} à deux méthodes d'apprentissage supervisé classiques, toutes deux basées sur des réseaux de neurones artificiels : le perceptron multi-couche (PMC) et les réseaux de fonctions à base radiale (FBR). Nous utilisons dans les deux cas des réseaux à une couche cachée et faisons varier le nombre de neurones de cette couche. Le PMC utilise la fonction sigmoïde comme fonction d'activation, le FBR utilise la fonction gaussienne. Pour évaluer les différentes méthodes supervisées, nous effectuons une validation croisée à 10 paquets (10 fold cross validation). Différentes mesures de qualité sont utilisées : précision, rappel, F-mesure, indice de Rand, indice de Jaccard. Ces mesures sont détaillées dans les équations suivantes, où vp , vn , fp et fn sont respectivement le nombre de vrais positifs, vrais négatifs, faux positifs et faux négatifs.

$$précision = \frac{vp}{vp + fp} \quad (15)$$

$$rappel = \frac{vp}{vp + fn} \quad (16)$$

$$F - mesure = 2 \cdot \frac{précision \cdot rappel}{précision + rappel} \quad (17)$$

$$indice\ de\ Rand = \frac{vp + vn}{vp + fp + vn + fn} \quad (18)$$

$$indice\ de\ Jaccard = \frac{vp}{vp + fp + fn} \quad (19)$$

2. University of Irvine California, www.ics.uci.edu/~mlearn/

Apprentissage de règles floues

dataset	#instances	#attributs	#classes
iris	150	4	3
column 2C	310	6	2
column 3C	310	6	3
wine	178	13	3
breast cancer	699	9	2
wdbc	569	30	2
yeast	1484	8	10
parkinson	195	22	2

TAB. 1 – Propriétés des différents jeux de données utilisés : nombre d’instances, nombre d’attributs et nombre de classes.

Les résultats obtenus sont présentés dans les tableaux 2 et 3. Nous pouvons voir que dans le cas non supervisé les deux algorithmes CAH et ARF_{CAH}^{auto} produisent des résultats médiocres dans l’ensemble. ARF_{CAH}^{auto} montre des résultats globalement meilleurs dans cinq jeux de données contre un seulement pour CAH. De plus, ARF_{CAH}^{auto} se démarque sur le jeu de données *breast cancer* avec une F-mesure supérieure à 0.9. Dans le cas supervisé les résultats sont très satisfaisants. Les mesures sont meilleures que celles obtenues avec ARF_{CAH}^{auto} . Les mesures sont supérieures à 0.8 dans six cas, voir supérieures à 0.9 dans deux cas. Notre algorithme est proche des résultats des approches basées sur des réseaux de neurones pour la majorité des jeux de données, sinon meilleurs pour trois jeux (*wine*, *wdbc* et *parkinson*).

dataset	méthode	P	R	FM	RI	JI
iris	CAH	0.436	0.492	0.462	0.624	0.301
	ARF_{CAH}^{auto}	0.625	0.683	0.653	0.761	0.484
column 2C	CAH	0.562	1	0.719	0.562	0.562
	ARF_{CAH}^{auto}	0.562	1	0.719	0.562	0.562
column 3C	CAH	0.374	1	0.544	0.374	0.374
	ARF_{CAH}^{auto}	0.471	0.697	0.563	0.595	0.391
wine	CAH	0.500	0.694	0.581	0.662	0.410
	ARF_{CAH}^{auto}	0.418	0.897	0.570	0.542	0.398
breast cancer	CAH	0.794	0.801	0.798	0.777	0.663
	ARF_{CAH}^{auto}	0.910	0.922	0.916	0.907	0.845
wdbc	CAH	0.678	0.745	0.710	0.676	0.550
	ARF_{CAH}^{auto}	0.702	0.865	0.775	0.733	0.633
yeast	CAH	0.251	0.544	0.344	0.537	0.208
	ARF_{CAH}^{auto}	0.271	0.326	0.296	0.655	0.174
parkinson	CAH	0.629	1	0.771	0.627	0.627
	ARF_{CAH}^{auto}	0.745	0.836	0.788	0.717	0.650

TAB. 2 – Résultats pour différents jeux de données. Deux méthodes non supervisées sont testées : CAH classique et ARF_{CAH}^{auto} . P=précision, R=rappel, FM=F-mesure, RI=indice de Rand, JI=indice de Jaccard.

dataset	méthode	P	R	FM	RI	JI
iris	ARF_{CAH}^{sup}	0.835	0.844	0.840	0.894	0.724
	PMC	0.973	0.974	0.973	0.982	0.948
	FBR	0.947	0.948	0.948	0.966	0.901
column 2C	ARF_{CAH}^{sup}	0.719	0.658	0.687	0.664	0.524
	PMC	0.562	1	0.719	0.562	0.562
	FBR	0.755	0.812	0.782	0.746	0.643
column 3C	ARF_{CAH}^{sup}	0.705	0.813	0.755	0.803	0.607
	PMC	0.385	0.844	0.528	0.437	0.359
	FBR	0.758	0.789	0.773	0.827	0.630
wine	ARF_{CAH}^{sup}	0.800	0.788	0.794	0.862	0.659
	PMC	0.351	0.611	0.445	0.486	0.287
	FBR	0.376	0.626	0.470	0.523	0.307
breast cancer	ARF_{CAH}^{sup}	0.905	0.857	0.880	0.872	0.786
	PMC	0.955	0.932	0.944	0.939	0.893
	FBR	0.953	0.944	0.949	0.944	0.903
wdbc	ARF_{CAH}^{sup}	0.732	0.690	0.710	0.701	0.551
	PMC	0.533	0.995	0.694	0.533	0.531
	FBR	0.603	0.665	0.632	0.589	0.462
yeast	ARF_{CAH}^{sup}	0.349	0.639	0.451	0.654	0.291
	PMC	0.425	0.460	0.441	0.741	0.283
	FBR	0.250	0.273	0.261	0.656	0.150
parkinson	ARF_{CAH}^{sup}	0.701	0.884	0.782	0.691	0.642
	PMC	0.627	1	0.771	0.627	0.627
	FBR	0.627	1	0.771	0.627	0.627

TAB. 3 – Résultats pour différents jeux de données. Trois méthodes supervisées sont testées : ARF_{CAH}^{sup} , le perceptron multi-couches et le réseau de fonctions à base radiale. P=précision, R=rappel, FM=F-mesure, RI=indice de Rand, JI=indice de Jaccard.

Il est important de remarquer ici que nous ne cherchons pas à obtenir des résultats parfaits mais à obtenir un algorithme simple et efficace qui donne des résultats cohérents dans un problème de faible taille (peu d'instances et peu de classes). L'objectif étant d'utiliser cette méthode dans un processus d'apprentissage plus complexe (voir section 4.1). En cela les résultats pour les jeux de données *iris*, *wine*, *breast cancer*, *wdbc* et *parkinson* sont satisfaisants. Ce choix a aussi été motivé par la rapidité de notre méthode dont le temps d'exécution est bien inférieur à la seconde. Notons aussi que PMC et FBR sont des approches classiques et ne sont pas utilisées ici pour apprendre des règles floues. Nous avons vu dans l'état de l'art de la section 3 que de nombreuses méthodes d'apprentissage de règles floues basées sur des réseaux de neurones existent et seraient peut-être plus adaptées dans notre cas. Cependant, ce n'est pas l'objectif des travaux présentés ici.

5 Expérimentations en télédétection

Dans cette expérience nous nous concentrons sur l'extraction de pavillons dans un quartier résidentiel. L'image utilisée est une image QUICKBIRD d'une résolution de 0,7 mètres, acquise en 2002, représentant un quartier de Toulouse. Comme énoncé dans la section 1, la segmentation d'une telle image ne peut que produire une sur-segmentation. Nous choisissons donc de considérer les pavillons comme des objets composés et notre but est d'apprendre une règle de composition décrivant le concept *pavillon*. Un bâtiment urbain étant en partie défini par son voisinage, nous choisissons d'utiliser la relation d'adjacence à la végétation et la route comme relations spatiales. La figure 6(a) montre l'image en entrée et les exemples fournis par l'expert. L'algorithme de segmentation utilisé est l'algorithme meanshift (Comaniciu et al., 2002), un exemple de segmentation est donné dans la figure 6(b). Après la segmentation, la végétation et les routes sont extraites. La végétation est extraite à l'aide de l'algorithme présenté dans la section 3.3, les routes sont extraites à l'aide du système d'information géographique (SIG).

Une fois l'image segmentée et transformée en graphe, le processus d'apprentissage peut prendre place : les exemples sont extraits, les règles sont apprises et les concepts correspondants sont générés puis utilisés pour étiqueter les objets. La figure 6(c) montre que deux concepts ont été appris (en jaune et bleu sur l'image). La végétation est représentée en vert, les routes en gris. Enfin, la règle de composition est apprise et le concept composé correspondant est généré. Le résultat est présenté dans la figure 6(d). Les pavillons sont extraits avec une précision de 0.82 après comparaison avec une vérité terrain fournie par l'expert géographe. Nous pouvons voir que certains pavillons pourtant évident pour un oeil humain n'ont pas été identifiés (par exemple les pavillons notés *a*, *b* et *c* sur l'image 6(d)). Après observation il apparait que l'objet composé correspondant au pavillon *a* est plus grand que les exemples fournis par l'expert et donc ne satisfait pas la contrainte de taille de la règle apprise, les segments composant le pavillon *b* sont mal identifiés dans la phase précédente, l'objet composé correspondant au pavillon *c* ne satisfait pas les contraintes de composition de la règle apprise. De plus il s'avère que la nature des pavillons n'est pas la même dans l'image : les îlots sont découpés différemment et contiennent des pavillons de différentes formes et tailles. Un biais est induit de par le fait que les exemples ne représentent pas l'intégralité des pavillons de l'image. Ceci montre que le choix et la qualité des exemples est crucial.

Des expériences supplémentaires ont été menées pour évaluer l'influence des paramètres d'apprentissage (les différentes valeurs du critère de Ward dans ce cas) ainsi que les paramètres de l'algorithme de segmentation. Les résultats montrent que les paramètres d'apprentissages influent peu alors que l'étape de segmentation est toujours critique. Les résultats présentés dans la figure 6 ont été obtenus avec les meilleurs paramètres trouvés dans cette analyse.

6 Conclusion

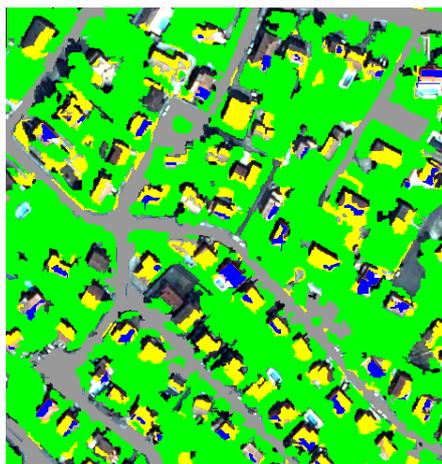
Dans cet article nous avons proposé une méthode d'apprentissage de règles floues multi-niveaux dans le but d'apprendre des relations de composition décrivant des objets d'intérêt. Les résultats présentés montrent l'intérêt de notre approche dans l'extraction d'objets composés à partir d'une image de télédétection à très haute résolution spatiale. Notre méthode a pour avantage de n'avoir besoin que de quelques exemples fournis par l'expert. De plus, notre méthode est indépendante du type d'image, de l'algorithme de segmentation utilisé ou de la



(a) Les exemples fournis par l'expert



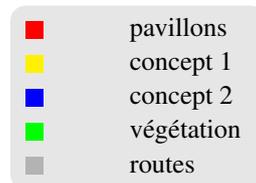
(b) Le résultat de la segmentation. En vert, les bords de chaque segment



(c) En bleu et jaune, les deux concepts appris par la méthode. La végétation en vert, les routes en gris



(d) Les pavillons extraits



(e) Légende

FIG. 6 – Application de la méthode sur une image QUICKBIRD (©DigitalGlobe 2001).

résolution, pour peu que les objets soient discernables sur l'image. Il est important de noter ici que notre méthode n'est pas un algorithme de détection de pavillons. Bien que les expériences aient été faites sur les pavillons, la méthode est générique et pourrait s'appliquer à tout type d'objet d'intérêt.

Nous avons vu dans la section 4.2 que la méthode ARF_{CAH}^{sup} est globalement efficace pour l'apprentissage de règles mais n'est pas optimale. Des travaux sont en cours pour améliorer l'apprentissage de règles floues. Notre intérêt se porte sur les méthodes basées sur les réseaux de neurones flous (vu dans l'introduction de la section 3) pour remplacer ARF_{CAH}^{sup} . Dans un second temps nous souhaitons aller plus loin dans l'utilisation de règles de composition et monter dans les niveaux d'abstraction pour extraire différentes couches d'objets composés (pavillons puis îlots urbains puis quartiers résidentiels, ...) dans le but de permettre une analyse globale d'un paysage urbain. Nous souhaitons de plus étudier la portabilité des règles apprises entre deux zones d'une même image ou entre deux images de même type (même capteur, même région, même période d'acquisition).

Références

- Adams, R. et L. Bischof (1994). Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(6), 641–647.
- Belarte, B. et C. Wemmert (2012). Apport des relations spatiales dans l'extraction automatique d'informations à partir d'image. In *Atelier Fouille de données complexes, Conférence EGC 2012*.
- Bilenko, M., S. Basu, et R. J. Mooney (2004). Integrating constraints and metric learning in semi-supervised clustering. In *Proceedings of 21st International Conference on Machine Learning (ICML-2004)*, Banff, Canada, pp. 81–88.
- Bittner, T. et S. Winter (1999). On ontology in image analysis. *International Workshop On Integrated Spatial Databases ISD'99 1737*, pp. 168–191.
- Castillo, L., A. González, et R. Pérez (2001). Including a simplicity criterion in the selection of the best rule in a genetic fuzzy learning algorithm. *Fuzzy Sets and Systems* 120(2), 309–321.
- Chakraborty, D. et N. Pal (2004). A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification. *IEEE Transactions on Neural Networks* 15(1), 110–123.
- Comaniciu, D., P. Meer, et S. Member (2002). Mean shift : A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 603–619.
- Cpalka, K. (2009). A new method for design and reduction of neuro-fuzzy classification systems. *IEEE Transactions on Neural Networks* 20(4), 701–714.
- Durand, N., S. Derivaux, G. Forestier, C. Wemmert, P. Gançarski, O. Boussaid D, et A. Puissant (2007). Ontology-based object recognition for remote sensing image interpretation. In *IEEE International Conference on Tools with Artificial Intelligence*, Patras, Greece, pp. 472–479.
- Fernández, A., M. Calderón, E. Barrenechea, H. Bustince, et F. Herrera (2010). Solving multi-class problems with linguistic fuzzy rule based classification systems based on pairwise learning and preference relations. *Fuzzy Sets and Systems* 161(23), 3064–3080.

- Fonseca, F., M. Egenhofer, P. Agouris, et G. Camara (2002). Using ontologies for integrated geographic information systems. *Transactions in Geographic Information Systems* 6(3), pp. 231–257.
- Forestier, G., S. Derivaux, C. Wemmert, et P. Gançarski (2008). An evolutionary approach for ontology driven image interpretation. In *Applications of Evolutionary Computing*, Volume 4974 of *Lecture Notes in Computer Science*, pp. 295–304. Springer Berlin Heidelberg.
- Forestier, G., A. Puissant, C. Wemmert, et P. Gançarski (2012). Knowledge-based region labeling for remote sensing image interpretation. *Computers, Environment and Urban Systems* 36(5), 470–480.
- Gruber, T. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies* 43(5/6), pp. 907–928.
- Gueguen, L., M. Pesaresi, A. Gerhardinger, et P. Soille (2012). Characterizing and counting roofless buildings in very high resolution optical images. *Geoscience and Remote Sensing Letters, IEEE* 9(1), 114–118.
- Ishibuchi, H. et T. Nakashima (2001). Effect of rule weights in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems* 9(4), 506–515.
- Ishibuchi, H. et T. Yamamoto (2004). Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. *Fuzzy Sets and Systems* 141(1), 59–88.
- Ke, C.-Q., L. Xia, G.-D. Tang, et X. Cao (2008). Detecting urban vegetation using a object-oriented method with quickbird imagery. Volume 3, pp. III–138–III–141.
- Kumar, N. et K. Kummamuru (2008). Semisupervised clustering with metric learning using relative comparisons. *Knowledge and Data Engineering, IEEE Transactions on* 20(4), pp. 496–503.
- Laha, A., N. Pal, et J. Das (2006). Land cover classification using fuzzy rules and aggregation of contextual information through evidence theory. *IEEE Transactions on Geoscience and Remote Sensing* 44(6), 1633–1641.
- Lance, G. N. et W. T. Williams (1967). A general theory of classificatory sorting strategies : 1. hierarchical systems. 9(4), 373–380.
- Lee, C. (1990a). Fuzzy logic in control systems : fuzzy logic controller. I. *Systems, Man and Cybernetics, IEEE Transactions on* 20(2), 404–418.
- Lee, C. (1990b). Fuzzy logic in control systems : fuzzy logic controller. II. *Systems, Man and Cybernetics, IEEE Transactions on* 20(2), 419–435.
- Maillot, N. E. et M. Thonnat (2008). Ontology based complex object recognition. *Image Vision Computing* 26(1), pp. 102–113.
- Mamdani, E. (1977). Application of fuzzy logic to approximate reasoning using linguistic synthesis. *Computers, IEEE Transactions on* C-26(12), 1182–1191.
- Mamdani, E. et S. Assilian (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* 7(1), 1–13.
- Mezaris, V., I. Kompatsiaris, et M. G. Strintzis (2004). Region-based image retrieval using an object ontology and relevance feedback. *EURASIP Journal on Advances in Signal Proces-*

- sing* 2004(1), pp. 886–901.
- Nakashima, T., G. Schaefer, Y. Yokota, et H. Ishibuchi (2007). A weighted fuzzy classifier and its application to image processing tasks. *Fuzzy Sets and Systems* 158(3), 284–294.
- Novák, V. (2012). Reasoning about mathematical fuzzy logic and its future. *Fuzzy Sets and Systems* 192(0), 25–44.
- Nozaki, K., H. Ishibuchi, et H. Tanaka (1996). Adaptive fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems* 4(3), 238–250.
- Pedrycz, W. (2004). Fuzzy clustering with a knowledge-based guidance. *Pattern Recognition Letters* 25(4), pp. 469–480.
- Russo, M. (2000). Genetic fuzzy learning. *IEEE Transactions on Evolutionary Computation* 4(3), 259–273.
- Sebari, I. et D.-C. He (2013). Automatic fuzzy object-based analysis of VHSR images for urban objects extraction. *ISPRS Journal of Photogrammetry and Remote Sensing* 79, 171–184.
- Suresh, S. et K. Subramanian (2011). A sequential learning algorithm for meta-cognitive neuro-fuzzy inference system for classification problems. In *The 2011 International Joint Conference on Neural Networks (IJCNN)*, pp. 2507–2512.
- Wagstaff, K., C. Cardie, S. Rogers, et S. Schroedl (2001). Constrained k-means clustering with background knowledge. In *Proceeding of the International Conference of Machine Learning*, pp. 577–584. Morgan Kaufmann.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58(301), 236–244.
- Wu, W., L. Li, J. Yang, et Y. Liu (2010). A modified gradient-based neuro-fuzzy learning algorithm and its convergence. *Information Sciences* 180(9), 1630–1642.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control* 8(3), 338–353.
- Zadeh, L. (1975). Fuzzy logic and approximate reasoning. *Synthese* 30(3), 407–428.

Summary

The new satellite sensors allow the acquisition of images of a very high level of detail at high speeds, thus producing a large amount of data. The manual processing of these data has become impossible, new tools are needed to process them automatically. In this context, effective segmentation algorithms are required to extract objects of interest of these images. However, the segments produced by the algorithm do not match the objects of interest. In this paper, we propose to change the level of abstraction in order to interpret the objects of interest as objects composed by segments. For this, we have implemented a multi-level learning process, based on expert knowledge, in order to learn the rules of compositions defining objects of interest. To manage the imprecision of the analysis of remote sensing images we propose to use fuzzy logic to model the rules of composition. The proposed method is validated on synthetic data and real data.