

# Segmentation sémantique de données de télédétection multimodale : application aux peuplements forestiers

## Semantic segmentation of multimodal remote sensing data : case study of forest stands

Clément Dechesne<sup>1</sup>

Clément Mallet<sup>1</sup>

Arnaud Le Bris<sup>1</sup>

Valérie Gouet-Brunet<sup>1</sup>

<sup>1</sup> Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, F-94160 Saint-Mande, France

73 avenue de Paris  
94165 Saint-Mandé  
clement.dechesne@ign.fr

### Résumé

*La délimitation des peuplements forestiers est une connaissance fondamentale pour la gestion des forêts, pour les politiques publiques, pour l'aménagement du territoire, etc. Cette tâche est encore principalement réalisée manuellement par photo-interprétation d'images géospatiales à (très) haute résolution spatiale (THR). La délimitation des peuplements a été peu abordée dans la littérature, et s'est principalement intéressée, dans des environnements forestiers, à l'extraction individuelle des arbres, qui reste imprécise, et la classification des espèces d'arbres, qui présente des résultats peu satisfaisants. Dans cet article, une méthode fondée sur la fusion des données lidar aéroportées et des images multispectrales THR est proposée pour la délimitation automatique des peuplements forestiers. Les images multispectrales donnent des informations sur les espèces d'arbres, tandis que les nuages de points lidar 3D fournissent des informations géométriques (sur la hauteur principalement). Des attributs multimodaux sont calculés, à la fois au niveau du pixel et de l'objet : les objets sont obtenus à partir d'une sur-segmentation. Une classification supervisée est ensuite effectuée au niveau de l'objet afin de discriminer grossièrement les espèces d'arbres existantes dans chaque zone d'intérêt. Les résultats de la classification sont ensuite traités pour obtenir des zones homogènes avec des frontières lisses par minimisation d'énergie. La formulation de l'énergie comporte deux parties ; une liée à la classification, et une autre dans laquelle des contraintes supplémentaires liées aux attributs précédemment calculés, sont ajoutées. Le modèle énergétique est efficace avec un gain de précision jusqu'à 15% par rapport à la classification au niveau de l'objet. Les résultats de la segmentation utilisant ce modèle ont une précision allant de 96% à 99% par rapport à la Base de Données Forêt de l'IGN.*

### Mots Clef

Lidar, imagerie multispectrale, fusion, sélection d'attributs, classification supervisée, minimisation d'énergie, régularisation, délimitation de peuplements forestiers.

### Abstract

*Forest stand delineation is a fundamental task for forest management purposes, for public policies, for land use planning, etc. It is still mainly manually performed through visual inspection of geospatial (very) high spatial resolution images (VHR). Stands detection has been barely addressed in the literature which has mainly focused, in forested environments, on individual tree extraction, that is still not very accurate, and tree species classification, that exhibits poor discrimination results. In this article, a method based on the fusion of airborne lidar data and VHR multispectral images is proposed for the automatic delineation of forest stands. The multispectral images give information about the tree species whereas 3D lidar point clouds provide geometric information (mainly on height). Multimodal features are computed, both at pixel and object levels : the objects are obtained from an over-segmentation. A supervised classification is then performed at the object level in order to coarsely discriminate the existing tree species in each area of interest. The classification results are further processed to obtain homogeneous areas with smooth borders by employing an energy minimum framework, where additional constraints related to the features computed previously, are joined to form the energy function. The energetical model is efficient with an accuracy gain up to 15% compared to the object-based classification. The segmentation results using such model has accuracy ranging from 96% to 99% compared to the IGN Forest Land Cover.*

### Keywords

Lidar, multispectral imagery, fusion, feature selection, supervised classification, energy minimisation, regularisation, forest stand delineation.

# 1 Introduction

L'extraction de l'information dans les zones forestières, en particulier au niveau du peuplement, est motivée par deux objectifs principaux : l'inventaire statistique et la cartographie. Les peuplements forestiers sont les unités de base et peuvent être définis en termes d'espèces d'arbres ou de maturité des arbres. Du point de vue de la télédétection, la délimitation des peuplements est un problème de segmentation sémantique. Pour l'inventaire forestier statistique, la segmentation est utile pour extraire des points d'échantillonnage significatifs sur le plan statistique et des attributs fiables (surface terrière, hauteur dominante, etc.) qui seront ensuite utilisés pour l'inventaire multi-source et pour la stratification [1, 2]. Pour la cartographie du couvert végétal, la segmentation est très utile pour la mise à jour des bases de données géographiques forestières [3]. La plupart du temps, pour des raisons de fiabilité, chaque zone est interprétée manuellement par des opérateurs humains avec des images géospatiales de très haute résolution spatiale ( $\leq 0.5$  m). En France, la photo-interprétation permet de produire la BD Forêt qui comporte la cartographie des peuplements forestiers. Ce travail prend beaucoup de temps. De plus, dans de nombreux pays, la grande variété d'essences forestières (environ 20 en France) rend le problème plus difficile.

L'utilisation de données de télédétection pour l'analyse automatique des forêts est de plus en plus répandue, en particulier avec l'utilisation combinée du lidar aéroporté et de l'imagerie optique (imagerie multispectrale à très haute résolution et imagerie hyperspectrale) [4].

Certains articles proposant une délimitation automatique des peuplements forestiers à partir de données de télédétection ont déjà été publiés. Tout d'abord, la délimitation peut être réalisée avec une seule source de télédétection (imagerie multispectrale/hyperspectrale ou lidar). Une technique de délimitation des peuplements utilisant des images hyperspectrales est proposée dans [5]. Les arbres sont extraits et classés selon 7 espèces (5 conifères, 1 caduque et 1 non spécifié) avec un classifieur par maximum de vraisemblance.

Une méthode de cartographie des peuplements utilisant des données lidar aéroportées à basse densité ( $< 6$  points/m<sup>2</sup>) est proposée dans [6]. Elle se compose de plusieurs étapes ; extraction d'attributs, *rastérisation* des attributs et classification à partir du raster. Les peuplements forestiers sont créés en regroupant les cellules voisines d'une même espèce au sein de chaque classe. Ensuite, seuls les peuplements ayant une taille minimale prédéfinie sont acceptés. Les petites zones voisines d'espèce différentes qui n'atteignent pas la taille minimale sont fusionnées à un peuplement forestier proche.

La délimitation des peuplements forestiers proposée dans [7] utilise aussi du lidar aéroporté à basse densité pour une segmentation et une classification supervisée. Trois attributs (couverture de la canopée, densité de la tige et hauteur moyenne) sont calculés et *rasterisés*. La segmentation est

réalisée par croissance de région. Les pixels spatialement adjacents sont regroupés en objets ou régions homogènes. Ensuite, une classification supervisée de l'image segmentée est réalisée à l'aide d'un classifieur utilisant la distance de Bhattacharya, les étiquettes correspondant à la maturité des peuplements.

Une délimitation de peuplements forestiers utilisant des données lidar aéroportées à haute densité est également proposée dans [8]. Trois attributs sont d'abord extraits du nuage de points ; un indicateur de la taille de l'arbre, un indice de densité forestière et un indicateur d'essences d'arbres. Une délimitation du peuplement forestier grossier est ensuite effectuée sur l'image des attributs en utilisant l'algorithme Mean-Shift, avec une valeur élevée des paramètres afin d'obtenir des peuplements forestiers grossiers sous-segmentés. Un masque forestier est ensuite appliqué à l'image segmentée afin de récupérer des peuplements forestiers et non-forestiers grossiers. Cette étape peut créer quelques petites zones isolées qui seront fusionnées à leurs voisins les plus proches jusqu'à ce que leur taille soit supérieure à un seuil défini par l'utilisateur. Les peuplements forestiers sont ensuite raffinés, mais, au lieu d'utiliser les pixels d'origine de l'image 3 bandes, des superpixels sont générés à partir des trois attributs. Le raffinement des peuplements forestiers est obtenu grâce à une croissance de région. Cette méthode fournit des peuplements relativement grands.

D'autres méthodes utilisant la fusion de différents types de données de télédétection ont également été développées. Deux méthodes de segmentation sont proposées dans [9] pour une forêt composée de pins sylvestres, d'épicéas de Norvège et de feuillus. La première est une segmentation sur la hauteur de couronne suivie d'une croissance de région itérative sur une image composite de lidar et d'image infrarouge. La deuxième méthode propose une segmentation hiérarchique sur la hauteur de couronne. Chaque objet de l'image est connecté à la fois aux objets de l'image de niveaux supérieur et inférieur. Cela permet de considérer les segments fins à partir des niveaux de segmentation les plus fins, comme les arbres individuels dans la zone.

L'analyse des données lidar et multispectrales est réalisée à trois niveaux dans [10]. Le premier niveau représente des petits objets (arbres individuels ou petits groupes d'arbres) qui peuvent être différenciés suivant des caractéristiques spectrales et structurelles en utilisant une classification fondée sur des règles. Le deuxième niveau correspond au peuplement. Il est construit en utilisant le même processus de classification que le niveau précédent, en se référant aux objets à petite échelle du niveau 1. Le troisième niveau est généré en fusionnant des objets du même développement forestier en unités spatiales plus grandes. Cette méthode produit une cartographie permettant d'évaluer la phase de développement forestier (les étiquettes ne correspondent pas aux espèces d'arbres).

Au regard des méthodes existantes, il semble qu'il n'y ait pas de méthode de segmentation des peuplements fores-

tiers, en termes d'espèce, capable de traiter de façon satisfaisante un grand nombre de classes ( $> 5$ ). Il apparaît également que le fait de travailler au niveau de l'objet (habituellement le niveau de l'arbre), afin de discriminer les essences d'arbres, produit de meilleurs résultats de segmentation des peuplements. Plusieurs méthodes de classification des essences au niveau des arbres ont été étudiées dans [5] et [11]. Toutefois, il est probable qu'elles soient imprécises, ce qui entraîne une classification inexacte des arbres. Cependant, le classement des essences forestières au niveau des arbres peut être utilisé pour la délimitation des peuplements forestiers.

Dans cet article, une méthode pour la segmentation des peuplements en termes d'espèce est proposée. La méthode est composée de trois étapes principales. Des attributs sont extraits au niveau du pixel et de l'objet. Les objets sont déterminés par une sur-segmentation. Une classification est effectuée au niveau de l'objet car elle améliore significativement les résultats de discrimination. Cette classification est ensuite régularisée par une minimisation d'énergie. La régularisation, réalisée avec une méthode de coupe de graphe, produit des zones homogènes en termes d'espèces d'arbres avec des frontières lisses. La fusion est effectuée à deux niveaux, celui des attributs, et lors de la classification et de la régularisation.

## 2 Méthode proposée

### 2.1 Extraction des attributs

L'extraction des attributs est composée de trois étapes :

- Calcul et *rastérisation* des attributs lidar.
- Calcul des attributs spectraux.
- Extraction des objets (sur-segmentation) et création des images avec des attributs à l'objet.

**Attributs lidar.** Les attributs lidar nécessitent de prendre en compte un voisinage pour être cohérent. Pour chaque point lidar, 3 voisinages cylindriques, avec le même axe vertical, sont utilisés (rayons de 1 m, 3 m et 5 m, hauteur infinie). Tout d'abord, deux attributs de densité de végétation sont calculés : le premier est fondé sur le nombre de maxima locaux dans les voisinages et le second est lié au nombre de points hors-sol dans le voisinage (les points au sol ayant été déterminés précédemment par filtrage).

Le scatter et la planarité sont aussi calculés en suivant la formulation proposée par [12]. Des attributs statistiques, reconnus pour être pertinents pour la classification des espèces [13, 14], sont aussi calculés. Pour chaque point lidar, les 3 mêmes voisinages cylindriques sont utilisés. Deux informations du lidar, la hauteur et l'intensité, sont utilisées afin de dériver des attributs statistiques. Un attribut statistique  $f_d$ , dérivé d'une information  $f_o$ , (hauteur ou intensité) est obtenu comme suit :

$$f_d = \frac{1}{3} \sum_{r \in \{1,3,5\}} f_s(\mathbf{p}_r, \mathbf{f}_o), \quad (1)$$

où  $f_s$  est une fonction statistique (minimum ; maximum ; moyenne ; médiane ; écart-type ; déviation absolue médiane d'une médiane ; déviation absolue moyenne d'une médiane ; *skewness* ; *kurtosis* ; 10<sup>ème</sup>, 20<sup>ème</sup>, 30<sup>ème</sup>, 40<sup>ème</sup>, 50<sup>ème</sup>, 60<sup>ème</sup>, 70<sup>ème</sup>, 80<sup>ème</sup>, 90<sup>ème</sup> et 95<sup>ème</sup> centile), et  $\mathbf{p}_r, \mathbf{f}_o$  un vecteur contenant les valeurs des points de l'information  $f_o$  dans le cylindre de rayon  $r$ . Toutes ces fonctions statistiques sont utilisées pour la hauteur. Seule la moyenne est utilisée pour l'intensité.

24 attributs sont calculés au cours de cette étape ; 2 liés à la densité de végétation, 2 liés à la distribution 3D locale du nuage de points et 20 attributs statistiques.

Les 24 attributs sont *rastérisés* à la même résolution spatiale que l'image multispectrale, en utilisant la méthode proposée dans [15]. Cette méthode est intéressante car elle produit des images lisses, qui permettent d'obtenir de meilleurs résultats pour la classification et la régularisation [16]. Le modèle numérique de surface (MNS) est aussi obtenu en utilisant cette méthode, à la même résolution spatiale, en utilisant un modèle numérique de terrain (MNT) précédemment obtenu en filtrant le nuage de points. Le MNS est très important car il permet de déterminer la hauteur par rapport au sol et est un attribut très discriminant pour la classification [17, 18]. Au total, 25 attributs lidar sont calculés.

**Attributs spectraux.** Les 4 bandes spectrales originales sont conservées et considérées comme des attributs spectraux. 3 indices de végétation pertinents ; le NDVI, le DVI et le RVI sont calculés car ils peuvent fournir plus d'information que les bandes originales seules. Comme pour les attributs lidar, des attributs statistiques sont calculés pour chaque bande et chaque indice de végétation en utilisant l'équation 1 (3 voisinages circulaires de rayon de 1 m, 3 m et 5 m). D'autres fonctions statistiques sont utilisées (minimum ; maximum ; moyenne ; médiane ; écart-type ; déviation absolue moyenne d'une médiane ; déviation absolue moyenne d'une moyenne ; déviation absolue médiane d'une médiane ; déviation absolue médiane d'une moyenne). Au total, 70 attributs spectraux sont calculés.

**Extraction d'objets.** L'extraction d'objets est une étape importante permettant d'améliorer les résultats de la classification. Cependant, les objets peuvent être extraits de différentes manières ; il est possible d'effectuer une simple sur-segmentation sur l'un des 95 attributs disponibles (Watershed), d'utiliser des algorithmes plus avancés (SLIC, Quickshift, etc.) [19] ou bien d'extraire les arbres directement à partir du nuage de points. Il est apparu que le choix de la méthode de sur-segmentation n'impacte que très peu les résultats finaux. Ici, la sur-segmentation est effectuée sur le MNS. Une fois la sur-segmentation effectuée, les attributs sont moyennés sur les objets.

### 2.2 Classification

La classification est composée de deux étapes ; tout d'abord, le nombre d'attributs est réduit, en sélectionnant uniquement les plus pertinents, puis la classification est ef-

fectuée sur les attributs sélectionnés.

**Sélection d'attributs.** À cause du grand nombre d'attributs disponibles, une étape de sélection d'attributs doit être mise en place. La sélection d'attributs comporte deux étapes ; la détermination du nombre d'attributs à sélectionner et la sélection des attributs.

L'algorithme *SFFS* (*Sequential Forward Floating Search*) [20] est utilisé pour les deux étapes. Cet algorithme présente deux avantages : (i) il peut être utilisé avec plusieurs scores de classification (dans cette étude, le coefficient Kappa), (ii) il permet d'accéder à l'évolution du score de classification en fonction du nombre d'attributs sélectionnés. L'algorithme SFFS sélectionne  $p$  attributs qui maximisent le score de sélection (ici, le coefficient Kappa).

**Classification.** Un classifieur supervisé est utilisé afin de discriminer les essences forestières fournies par une base de données de couverture forestière existante (BD Forêt). La classification est réalisée par forêts aléatoires. Cette méthode de classification a démontré sa pertinence dans la littérature [21] et dans une étude précédente comparée à SVM [22], car elle fournit des résultats similaires tout en étant plus rapide. Les résultats de la classification sont (i) la carte de labels et (ii) la carte de probabilité (probabilités à posteriori de classe pour chaque pixel). Cette carte de probabilité est nécessaire pour l'étape de régularisation suivante.

### 2.3 Régularisation

La classification peut ne pas être suffisante pour obtenir des zones homogènes avec des frontières lisses. Par conséquent, afin de s'adapter au modèle de peuplement, la régularisation de classification au niveau des pixels par minimisation de l'énergie est une solution fiable. De plus, elle permet d'insérer des contraintes supplémentaires pour une délimitation plus souple des peuplements forestiers.

**Formulation de l'énergie.** Le modèle énergétique proposé repose sur un modèle graphique, c'est-à-dire que le problème est modélisé par un graphe probabiliste prenant en compte les probabilités de classe  $P$  et la carte d'attributs normalisés (moyenne nulle et écart-type unitaire)  $A$ . Il se compose d'un terme d'attache à la donnée et d'un terme de régularisation (*pairwise*). Pour une image  $I$  et une classification  $C$ , l'énergie  $E$  est formulée comme suit :

$$E(I, C) = \sum_{u \in I} E_{\text{data}}(C(u)) + \gamma \sum_{u, v \in \mathcal{N}} E_{\text{regul}}(C(u), C(v)), \quad (2)$$

$$\begin{aligned} \text{avec :} \quad E_{\text{data}}(C(u)) &= f(P(C(u))); \\ E_{\text{regul}}(C(u) = C(v)) &= 0; \\ E_{\text{regul}}(C(u) \neq C(v)) &= V(u, v). \end{aligned}$$

$\mathcal{N}$  est le voisinage en 8-connexité, et  $P(C(u))$  est la probabilité qu'un pixel  $u$  appartienne à la classe  $C$ .  $\gamma$  est le paramètre de régularisation permettant de contrôler le niveau d'homogénéité des segments ( $\gamma \in [0, \infty[$ ).

$E_{\text{data}}$  est le terme d'attache aux données.  $E_{\text{regul}}$  est un terme

permettant de mesurer la différence entre les attributs du pixel  $u$  et les attributs de ses 8 voisins. L'énergie exprime à quel point les pixels sont bien classés et à quel point les attributs sont similaires.

D'autres modèles de champs aléatoires conditionnels peuvent être envisagés pour l'expression de l'énergie [23, 24, 25].  $E_{\text{regul}}$  pourrait être exprimée relativement à plus de pixels, mais seul le voisinage en 8-connexité est considéré ici.

La fonction  $f$  liée à  $E_{\text{data}}$  (voir Équation (2)) est :

$$f(x) = 1 - x, \quad \text{avec } x \in [0, 1]. \quad (3)$$

Cette fonction permet de contrôler l'importance à donner à la classification. Une simple fonction linéaire permet de contrôler l'énergie : quand la probabilité est proche de zéro, l'énergie est maximale. Inversement, quand la probabilité est forte, l'énergie est minimale. Cette fonction permet de conserver ce terme dans  $[0, 1]$  et ainsi simplifier le paramétrage de  $\gamma$ .

Le terme de régularisation est une mesure qui contrôle la valeur de l'énergie en fonction de la valeur des attributs des 8 voisins. Deux pixels de classes différentes, mais avec des valeurs d'attributs proches sont plus susceptibles d'appartenir à la même classe que deux pixels de classes différentes avec des valeurs d'attributs différentes. La valeur de l'énergie doit être proche de 1 lorsque les valeurs des attributs sont assez similaires et diminue quand elles sont différentes. 3 formulations ont été testées pour le terme de régularisation  $V$  :

$$V(u, v) = 1, \quad (4)$$

$$V(u, v) = \frac{1}{n} \sum_{i=1}^n \exp(-\lambda_i |A_i(u) - A_i(v)|), \quad (5)$$

$$V(u, v) = 1 - \|A(\mathbf{u}); A(\mathbf{v})\|_{n,2}. \quad (6)$$

Le terme (4) correspond à un *modèle de Potts* classique ; la pénalisation est la même pour tous les voisins appartenant à une classe différente.

Le terme (5) est un *modèle sensible aux valeurs des attributs* (modèle exponentiel), dans lequel un poids unique peut être attribué à chaque attribut ;  $\forall i \in [1, n], \lambda_i \in [0, \infty[$ , où  $i$  est l'indice correspondant à l'attribut.  $A_i(u)$  est la valeur du  $i^{\text{ème}}$  attribut du pixel  $u$ ,  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_n]$  est un vecteur de longueur égale au nombre d'attributs ( $n$ ). Ce vecteur assigne des poids différents aux attributs. Si  $\lambda_i = 0$  l'attribut  $i$  ne sera pas pris en compte dans le processus de régularisation. Quand  $\lambda_i$  est important, une petite différence entre les attributs impactera beaucoup l'énergie. Ainsi, l'attribut  $i$  aura un fort impact dans la régularisation. Dans cet article,  $\lambda_i$  est fixé à 1.

Le terme (6) est un *modèle sensible à la distance entre les attributs* (modèle de distance) qui nécessite de calculer la valeur  $\|A(\mathbf{u}); A(\mathbf{v})\|_{n,2} = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^n (A_i(\mathbf{u}) - A_i(\mathbf{v}))^2}$ . Dans cette formulation aussi, quand deux pixels ont des attributs proches, la pénalisation est forte et décroît à mesure de leur différence. Les

attributs étant de différents types (hauteur, réflectance, densité, etc.), il est donc important de garder les termes dans  $[0, 1]$  pour chaque attribut, même s'ils ne sont pas dans la même gamme de valeurs, afin de s'assurer que l'énergie ne sera pas négative.

**Minimisation d'énergie.** La minimisation de l'énergie est réalisée en utilisant une méthode de coupe de graphe. L'algorithme de coupe de graphe utilisé est l'optimisation pseudo-booléenne quadratique (QPBO) avec  $\alpha$ -expansion. Le QPBO est une méthode de coupe de graphe très répandue qui résout efficacement les problèmes de minimisation d'énergie en construisant un graphe et en calculant la coupe minimale [26]. L' $\alpha$ -expansion permet de traiter les problèmes multi-classes [27].

Afin de permettre le passage à l'échelle, le graphe est calculé et résolu sur une fenêtre glissante  $N \times N$  ( $N = 1400$  dans nos expériences) et seuls les pixels centraux de la fenêtre  $l \times l$  ( $l = 500$  dans nos expériences) sont conservés. Le processus est itéré jusqu'à ce que toute la zone soit traitée. Les expériences ont montré que le choix de  $N$  et  $l$  n'a pas d'impact sur les résultats ( $N$  doit être supérieur à  $l$ ), il ne vise qu'à réduire la charge de calcul en résolvant un plus petit graphe.

Lorsque  $\gamma = 0$ , le terme de régularisation n'a aucun effet dans la formulation énergétique ; La classe la plus probable est attribuée au pixel, conduisant au même résultat que la sortie de classification. Lorsque  $\gamma \neq 0$ , la carte d'étiquette résultante devient plus homogène, et les bords des segments sont plus lisses. Cependant, si  $\gamma$  est trop élevé, les petites zones sont liées pour être fusionnées dans des zones plus grandes, en supprimant une partie des informations utiles fournies par l'étape de classification.

### 3 Données

Les zones d'étude se situent sur des forêts de montagne. Chacune des 4 zones comprend un grand nombre d'espèces (4-5 espèces par zone), permettant de tester la robustesse de la méthode à un grand nombre de classes.

Les images multispectrales aéroportées ont été acquises par les caméras numériques IGN. Elles sont composées de 4 bandes : 430-550nm (bleu), 490-610nm (vert), 600-720nm (rouge) et 750-950 nm (proche infrarouge) avec une résolution spatiale de 0,5 m.

Les données lidar aéroportées ont été recueillies en utilisant un dispositif Optech 3100EA. L'empreinte était de 0,8 m afin d'augmenter la probabilité d'atteindre le sol. La densité de points pour tous les échos varie de 2 à 4 points/m<sup>2</sup>. Nos données multispectrales et lidar sont en adéquation avec les normes utilisées dans de nombreux pays pour des fins de cartographie forestière à grande échelle [28]. Le recalage entre le nuage de points lidar et les images multispectrales VHR a été réalisé par l'IGN par un processus de production standard et maîtrisé.

La BD Forêt est composée de polygones 2D délimités par photo-interprétation. C'est une donnée nationale française pour les forêts, librement accessible. Elle est utilisée pour

entraîner le classifieur et pour évaluer les résultats. Seuls les polygones "purs", contenant au moins 75% d'une espèce donnée, sont utilisés pour la classification. Comme cette étude n'est basée que sur des espèces, la vérité terrain utilisée ne couvrira donc pas l'ensemble de la zone (les zones contenant du mélange ne seront pas prise en compte).

## 4 Résultats

### 4.1 Sélection d'attributs

L'algorithme SFFS permet de déterminer le nombre optimal d'attributs à sélectionner, et de les sélectionner. Le nombre optimal d'attributs à sélectionner est de 20. Ce nombre a été conservé pour les autres zones et les résultats montrent que le transfert du choix du nombre d'attributs n'a aucun impact sur la précision finale. Une fois le nombre optimal d'attributs déterminé, la sélection est effectuée 40 fois sur toutes les zones de test afin de récupérer les attributs les plus pertinents (les plus souvent sélectionnés). En moyenne, 61% des attributs sélectionnés proviennent de l'information spectrale et 39% de l'information lidar. Ceci montre la complémentarité des deux données de télédétection.

### 4.2 Classification

Les résultats de la classification et leur impact sur la segmentation finale sont illustrés dans la figure 2 pour une seule zone. La précision globale est obtenue en comparant chaque pixel de la BD Forêt aux résultats de classification. Il apparaît clairement que l'approche basée sur les pixels conduit à des cartes d'étiquettes bruitées, même si l'on obtient des résultats corrects (75% de précision globale). D'un autre côté, la classification sur les objets produit des étiquettes plus cohérentes sur le plan spatial (81,4% de précision globale). Cette différence est directement reflétée sur la segmentation finale. En effet, la régularisation de la classification pixellaire a une précision globale de 94,6%, tandis que la régularisation de la classification objets a une précision globale de 96,4%. Ces résultats confirment la pertinence du niveau objet pour la classification. Des résultats similaires ont également été observés pour les autres zones.

### 4.3 Régularisation

Les résultats finaux sont présentés dans la Figure 1. La précision globale montre que la méthode donne des résultats satisfaisants en termes de discrimination des espèces d'arbres. Les résultats sont améliorés jusqu'à 15% après régularisation. Les résultats doivent cependant être considérés avec précaution. Ils sont comparés à une BD qui comporte des défauts et ne couvre pas l'intégralité des zones d'étude. Après une analyse visuelle, il apparaît que les scores de segmentation peuvent en fait être pris avec une confiance de  $\pm 5\%$ . La table 1 présente les résultats des différentes méthodes de régularisation. Les trois méthodes présentent des résultats similaires, cependant, le modèle exponentiel présente les meilleurs résultats et conserve une

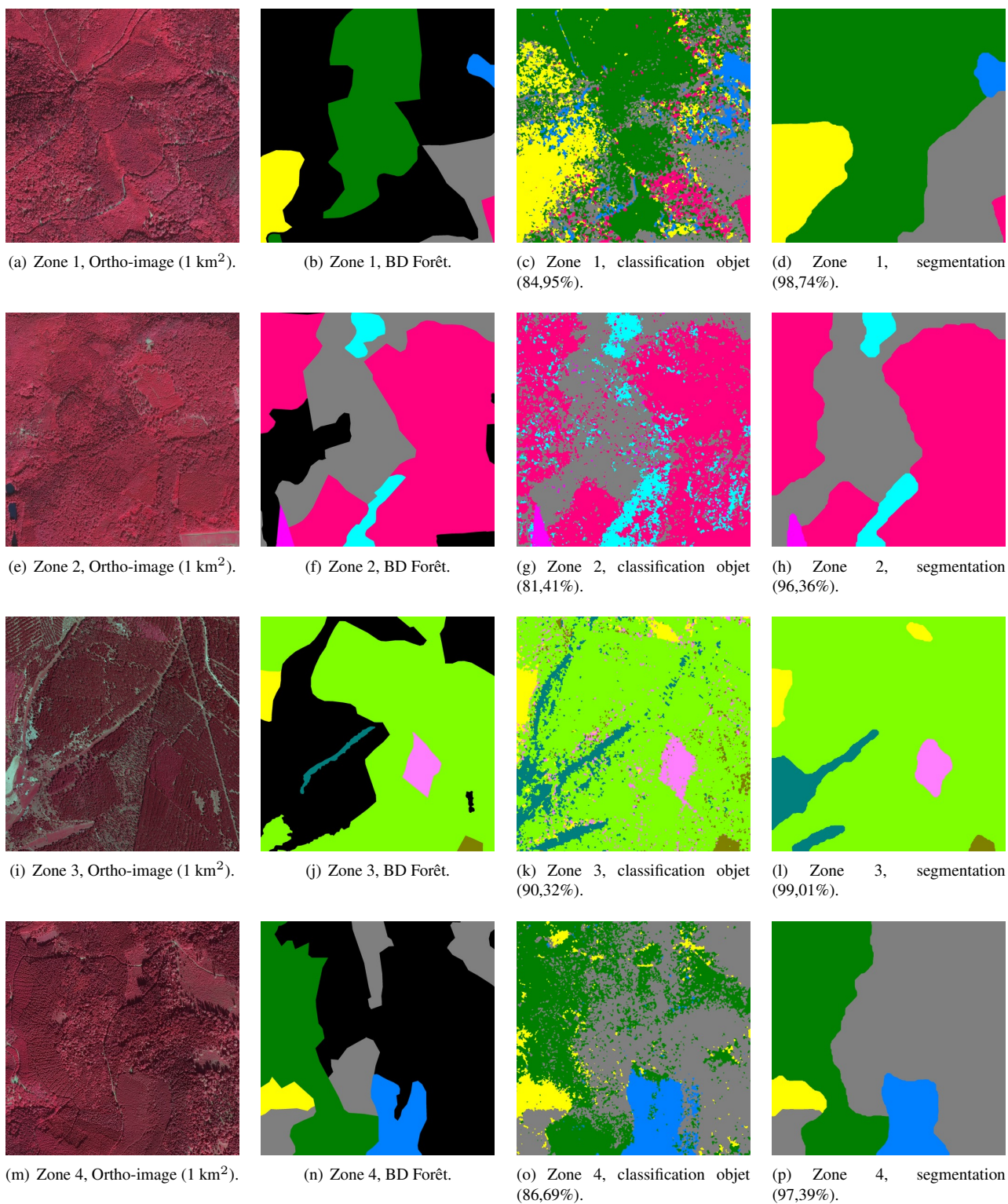


FIGURE 1. Résultats pour les zones testées, la précision globale est précisée entre parenthèse). La régularisation a été effectuée pour  $\gamma = 10$ . Code couleur : ● hêtre, ● chênes décidus, ● pin sylvestre, ● douglas, ● sapin ou épicéa, ● châtaigner, ● robinier, ● mélèze, ● sapin non pectiné, ● pin noir, ● formation herbacée, ● pas de données.

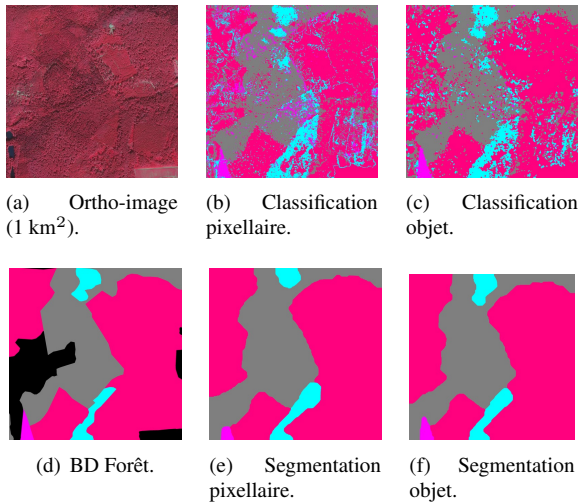


FIGURE 2. Résultats de classification et de régularisation (zone 2); comparaison entre l’approche pixellaire et l’approche objet. Code couleur : ● *chênes décidus*, ● *sapin ou épicéa*, ● *châtaigner*, ● *robinier*, ● pas de données.

bonne précision même quand le paramètre  $\gamma$  est mal choisi. Le réglage de  $\gamma$  est également une étape importante ; quand il est trop bas (par exemple  $< 5$ ), certains petits segments non pertinents peuvent rester dans la segmentation finale, ce qui entraîne un score de segmentation faible. Cependant, cette sur-segmentation peut être utile pour l’inventaire forestier statistique et la cartographie, car elle fournit de petites zones homogènes au sein de grands peuplements. Lorsque  $\gamma$  se trouve dans  $[10, 15]$ , la précision de segmentation est élevée : chaque classe est correctement étiquetée à plus de 95%. Cependant, avoir un  $\gamma$  élevé ( $> 20$ ) diminue légèrement la précision de la segmentation, mais réduit également considérablement la précision des classes sous-représentées (une classe ayant une petite superficie sera fusionnée avec une classe ayant une plus grande superficie).

Modèle	Précision de classification	Précision de régularisation	$\gamma$
Potts	81,41%	95,96%	5
		96,24%	10
		94,08%	15
		92,32%	20
Exponentiel	81,41%	95,6%	5
		<b>96,36%</b>	10
		95,27%	15
		94,09%	20
Distance	81,41%	95,63%	5
		96,23%	10
		94,12%	15
		92,34%	20

TABLEAU 1. Résultats des différentes méthodes de régularisation proposées (zone 2).

## 5 Conclusion

Une méthode en trois étapes pour la délimitation des peuplements forestiers, en termes d’espèce, a été proposée. La fusion des données lidar aéroporté et des images multispectrales à plusieurs niveaux produit des résultats très satisfaisants puisque les deux modalités de télédétection fournissent des observations complémentaires. Les segments finaux ont un bon score de correspondance avec les peuplements délimités par les opérateurs humains. La méthode repose sur le calcul d’attributs lidar et spectraux à différents niveaux (pixel et objet) pour une classification supervisée des espèces d’arbres. De bons scores de classification sont obtenus, produisant une base solide pour une délimitation des peuplements. Cette délimitation est alors réalisée grâce à une régularisation. Elle repose sur un modèle énergétique formulé en fonction des résultats des probabilités de classification et des valeurs des attributs. Elle permet d’obtenir des zones homogènes en termes d’espèces avec des frontières lisses. De plus, il est possible de contrôler le niveau des détails requis pour la délimitation du peuplement, qui dépend de l’inventaire forestier ou des spécifications de la BD.

L’objectif étant de délimiter les peuplements forestiers en fonction du type de végétation (principalement des espèces d’arbres), l’utilisation d’images hyperspectrales pourrait être intéressante pour obtenir plus d’informations sur les espèces. D’autres indices de végétation peuvent également être dérivés à partir données hyperspectrales. De plus, l’utilisation de données lidar à haute densité ( $\sim 10$  pts/m<sup>2</sup>) permettrait d’extraire les arbres de façon plus précise et de dériver de nouveaux attributs.

## Références

- [1] J.E. Means, S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix, “Predicting forest stand characteristics with airborne scanning lidar,” *Photogrammetric Engineering & Remote Sensing*, vol. 66, no. 11, pp. 1367–1372, 2000.
- [2] A. Kangas and M. Maltamo, *Forest inventory : methodology and applications*, vol. 10, Springer Science & Business Media, 2006.
- [3] M. Kim, M. Madden, and T.A. Warner, “Forest type mapping using object-specific texture measures from multispectral Ikonos imagery : Segmentation quality and image classification issues,” *Photogrammetric Engineering & Remote Sensing*, vol. 75, no. 7, pp. 819–829, 2009.
- [4] H. Torabzadeh, F. Morsdorf, and M. E. Schaepman, “Fusion of imaging spectroscopy and airborne laser scanning data for characterization of forest ecosystems—a review,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 97, pp. 25–35, 2014.
- [5] D. G Leckie, F. A Gougeon, N. Walsworth, and D. Paradine, “Stand delineation and composition estima-

- tion using semi-automated individual tree crown analysis,” *Remote Sensing of Environment*, vol. 85, no. 3, pp. 355–369, 2003.
- [6] B. Koch, C. Straub, M. Dees, Y. Wang, and H. Weinacker, “Airborne laser data for stand delineation and information extraction,” *International Journal of Remote Sensing*, vol. 30, no. 4, pp. 935–963, 2009.
- [7] A. A. Sullivan, R. J. McGaughey, H.-E. Andersen, and P. Schiess, “Object-oriented classification of forest structure from light detection and ranging data for stand mapping,” *Western Journal of Applied Forestry*, vol. 24, no. 4, pp. 198–204, 2009.
- [8] Z. Wu, V. Heikkinen, M. Hauta-Kasari, J. Parkkinen, and T. Tokola, “ALS data based forest stand delineation with a coarse-to-fine segmentation approach,” in *IEEE CISP*, Dalian, China, 2014, pp. 547–552.
- [9] V.J. Leppänen, T. Tokola, M. Maltamo, L. Mehtälä, T. Pusa, and J. Mustonen, “Automatic delineation of forest stands from lidar data,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38(4/C1)*, pp. 5–8, 2008.
- [10] D. Tiede, T. Blaschke, and M. Heurich, “Object-based semi automatic mapping of forest stands with laser scanner and multi-spectral data,” *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36(8/W2), pp. 328–333, 2004.
- [11] M. Dalponte, L. T. Ene, C. Witte, M. Marconcini, T. Gobakken, and E. Næsset, “Semi-supervised SVM for individual tree crown species classification,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 110, pp. 77–87, 2015.
- [12] M. Weinmann, B. Jutzi, S. Hinz, and C. Mallet, “Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 105, pp. 286 – 304, 2015.
- [13] M. Dalponte, H. O. Ørka, L. T. Ene, T. Gobakken, and E. Næsset, “Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data,” *Remote Sensing of Environment*, vol. 140, pp. 306–317, 2014.
- [14] H. Torabzadeh, R. Leiterer, M.E. Schaepman, and F. Morsdorf, “Optimal structural and spectral features for tree species classification using combined airborne laser scanning and hyperspectral data,” in *IGARSS, 2015 IEEE International*, Milano, Italy, 2015, IEEE, pp. 5399–5402.
- [15] A. Khosravipour, A. K. Skidmore, M. Isenburg, T. Wang, and Y. A. Hussin, “Generating pit-free canopy height models from airborne lidar,” *Photogrammetric Engineering & Remote Sensing*, vol. 80, no. 9, pp. 863–872, 2014.
- [16] J. Li, B. Hu, and T. L. Noland, “Classification of tree species based on structural features derived from high density lidar data,” *Agricultural and Forest Meteorology*, vol. 171–172, pp. 104 – 114, 2013.
- [17] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, “Relevance assessment of full-waveform lidar data for urban area classification,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. S71–S84, 2011.
- [18] M. Weinmann, “Reconstruction and analysis of 3d scenes,” *Springer International Publishing, Cham. doi*, vol. 10, pp. 978–3, 2016.
- [19] David Stutz, Alexander Hermans, and Bastian Leibe, “Superpixels : An evaluation of the state-of-the-art,” *arXiv preprint arXiv :1612.01601*, 2016.
- [20] P. Pudil, J. Novovičová, and J. Kittler, “Floating search methods in feature selection,” *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119–1125, 1994.
- [21] M. Belgiu and L. Drăguț, “Random forest in remote sensing : A review of applications and future directions,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, 2016.
- [22] Clément Dechesne, Clément Mallet, Arnaud Le Bris, and Valérie Gouet-Brunet, “Semantic segmentation of forest stands of pure species combining airborne lidar data and very high resolution multispectral imagery,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 126, pp. 129–145, 2017.
- [23] K. Schindler, “An overview and comparison of smooth labeling methods for land-cover classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4534–4545, 2012.
- [24] M. Volpi and V. Ferrari, “Semantic segmentation of urban scenes by learning local class interactions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Boston, MA, USA, 2015, pp. 1–9.
- [25] D. Tuia, M. Volpi, and G. Moser, “Getting pixels and regions to agree with conditional random fields,” in *IGARSS, 2016 IEEE International*, Beijing, China, 2016, IEEE, pp. 5399–5402.
- [26] V. Kolmogorov and C. Rother, “Minimizing non-submodular functions with graph cuts—a review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1274–1279, 2007.
- [27] V. Kolmogorov and R. Zabih, “What energy functions can be minimized via graph cuts?,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.
- [28] J. C. White, N. C Coops, M. A. Wulder, M. Vastaranta, T. Hilker, and P. Tompalski, “Remote sensing technologies for enhancing forest inventories : A review,” *Canadian Journal of Remote Sensing*, vol. 42, no. 5, pp. 619–641, 2016.