HAL
open science

Contributions méthodologiques pour la caractérisation des milieux par imagerie optique et LiDAR

Nesrine Chehata

To cite this version:

Nesrine Chehata. Contributions méthodologiques pour la caractérisation des milieux par imagerie optique et LiDAR. Apprentissage [cs.LG]. université paris-est, 2017. tel-01494206

HAL Id: tel-01494206

https://hal.science/tel-01494206

Submitted on 22 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mémoire d'habilitation à diriger des recherches

Spécialité Informatique

Nesrine Chehata
Maître de conférences en Informatique

Contributions méthodologiques pour la caractérisation des milieux par imagerie optique et LiDAR

Habilitation à diriger les recherches soutenue le 20 Mars 2017 devant le jury composé de :

Pr.	Jocelyn Chanussot	INP Grenoble	(Rapporteur)
Pr.	Florence Tupin	Télécom Paris	(Rapporteur)
Pr..	Laurent Wendling	Université Paris 5	(Rapporteur)
Pr.	Pierre Gançarski	Université de Strasbourg	(Examinateur)
Pr.	Christian Heipke	IPI, Leibniz Universität, Hannover	(Examinateur)
Pr.	Nicolas Paparoditis	Directeur de la recherche, IGN	(Garant HDR)

À Habib

Remerciements

Je voudrais tout d'abord remercier Pr. Jocelyn Chanussot, Pr. Laurent Wendling et Pr. Florence Tupin d'avoir accepté d'être rapporteur de mon manuscrit d'habilitation à diriger des recherches. Je remercie également Pr. Pierre Gançarski, Pr. Christian Heipke et Dr. Nicolas Paparoditis d'avoir accepté de faire partie de mon jury d'HDR.

Je souhaite ensuite exprimer mes plus profonds remerciements à tous les êtres qui sont chers à mon cœur et qui ont été à mes côtés au quotidien, m'ont appuyée et motivée pour rédiger mon habilitation ; mon cher mari Maher, mon fils Habib, mes parents sans lesquels ce travail n'aurait pas pu aboutir, et mon frère Amine.

Durant ce bilan de carrière, je me suis aussi remémoré tous les moments clés passés avec mes collègues de travail que je cite par ordre chronologique :

Mes collègues de l'équipe SIP du laboratoire LIPADE de l'université Paris 5 qui m'ont accueillie en thèse : à commencer par mon mentor Pr. Georges Stamon, mon père spirituel, à qui je souhaite une bonne santé et une longue vie, Pr. Nicole Vincent pour son appui, Florence Cloppet et Nicolas Loménie.

Un énorme merci à tous les membres du laboratoire MATIS de l'IGN (devenue maintenant équipe du laboratoire LaSTIG), à travers les années, les actions de recherche DATCAL, ARCHI, RISOTO. Avec mes mobilités fréquentes entre Paris, Bordeaux, Tunis et Montpellier, vous avez été mon repère fixe, vous avez toujours été là pour m'appuyer dans mes recherches et dans les avancées méthodologiques. La liste est longue et j'en oublierai sûrement, je remercie Marc Pierrot Desselligny de m'avoir accueillie en thèse au MATIS, les différents directeurs du laboratoire MATIS : Didier Boldo, Nicolas Paparoditis, Valérie Gouet, merci d'avoir cru en moi et d'avoir maintenu nos collaborations étroites contre vents et marées. Un grand merci à mes collègues des différents bureaux Frédéric Bretar, Franck Jung, Franck Taillandier, Bahman Soheilian, Olivier Tournaire, Olivier De Joinville qui sont partis de l'IGN, et ceux que je croise encore dans les couloirs de l'IGN : Grégoire Maillet, Franck Fuchs, François Boyero... L'équipe des "Hyènes" avec le grand chef Clément Mallet, Nicolas David, Arnaud Le-Bris, Antonio Ferraz, et la dernière recrue Sébastien Giordano. Un grand merci à Nicolas David et Arnaud Le-Bris pour tous nos échanges scientifiques et pour la relecture de ce manuscrit. De très bons souvenirs passés avec vous tous.

Mes remerciements les plus vifs vont ensuite à mes collègues bordelais que je connais depuis une dizaine d'année maintenant, au sein de l'institut EGID devenu ENSEGID, une équipe à taille humaine, une ambiance familiale, merci de m'avoir permis de choisir ma voie. Merci tout particulièrement à mes

amis les plus proches Myriam Schmutz, Amira Meddour, Michel Franceschi et François Larroque qui m'ont toujours soutenue. Merci à Jean Marie Malézieux (ex directeur EGID) et Philippe Paillou, professeur à l' université de Bordeaux et ex-EGID d'avoir été une locomotive pour m'aider à trouver ma place au sein d'une équipe en géosciences et pour tous leurs conseils et leur appui. Merci à tous mes partenaires bordelais, avec lesquels j'ai collaboré sur des aspects thématiques : Dominique Guyon, Jean Pierre Wigneron,Thierry Belouard, Christian Germain, Virginie Lafon, Aurélie Dehouck...

Un grand remerciement à toutes les personnes que j'ai côtoyées à l'UMR LISAH, durant ma mise en délégation à l'IRD Tunis pour deux ans. Mon passage au LISAH a été très enrichissant pour moi et a constitué le début de ma conversion thématique vers les ressources en eau et l'agriculture. Thématique à fort enjeu dans mon pays natal, la Tunisie et dans laquelle je souhaite vivement m'investir dans les prochaines années. Un grand merci à mon collègue de bureau Denis Feurer pour toutes nos discussions diverses et variées, à Zakia Jenhaoui pour son expertise sur le terrain et ses leçons de vie au quotidien, à Philippe Lagacherie et Jean Stéphane Bailly en tant que chefs de l'équipe paysage de m'avoir orientée et appuyée pour mes recherches passées et actuelles, à l'équipe de direction Jérôme Molénat et Olivier Grunberger de m'avoir également appuyée. Un grand merci à Frédéric Jacob, représentant de l'UMR LISAH à Tunis de 2012 à 2016 pour tous ses conseils scientifiques et administratifs ! Je suis sortie grâce à lui du statut d'"OVNI administratif".

Merci à mes collègues tunisiens avec lesquels j'ai collaboré en enseignement et en recherche, ou encore à travers notre association AGEOS (African Association for Geospatial Development), Pr. Hedia Chakroun, Pr. Riadh Abdelfattah, Pr. Rached Boussema, Riadh Tebourbi, Pr. Amel Ben Azza, Insaf Mekki, Thouraya Sahli, Pr. Riadh Imed Farah, Faten Chaieb, Pr. Zohra Lili Chabaane, Zeineb Kassouk, Karim Boudabous, Cheher Hamzaoui, Mohmad Chouchane, Omar Gaafar, Salwa Saidi, Pr. Beya Tayachi, Pr. AbdelAziz Kallel, Pr. Kamel Besbes,...

Merci à toutes ses personnes que je n'ai jamais voulu quittées professionnellement et avec lesquels je continue de collaborer et nombre d'entre eux sont devenus des amis.

Je n'oublie pas tous mes étudiants que j'ai eu la chance de côtoyer et de suivre, et tout particulièrement, les promos 2007-2011 du parcours géoimagerie de l'EGID, qui ont presque tous poursuivi dans le domaine de la télédétection et en recherche ; Cyrielle Guérin, Benoit Beguet, Victoria Meyer, Fabrice Monnier, Elvis Tshibasu, Nejeh Sayah, Ahmed Ezzine, Driss Bru, Olivier Regniers... Merci à tous les étudiants que j'ai encadrés en stage de PFE, en master ou en thèse Camille, Mahmoud, Karim, Walid, Safa, Houssem, Li, Benoit, Arnaud, Ines,... et sans lesquels je n'aurai pas pu avancer mes recherches. Un grand merci à vous tous, le travail avec vous me confirme chaque jour, toute la grandeur du métier d'enseignant. Je vous souhaite à tous le meilleur.

Je finirai par une spéciale dédicace à mon petit Habib, devenu grand du haut de ses 5 ans, mon rayon de soleil au quotidien qui me donne la force d'avancer et de donner le meilleur de moi-même.

Table des matières

Table des matières vii

Liste des figures ix

Avant-propos 1

1 Parcours et curriculum vitae 3
1.1 parcours . . . 5
1.2 Curriculum Vitae . . . 8

2 Besoins, données et enjeux scientifiques 31
2.1 Besoins en cartographie de l'occupation du sol . . . 33
2.1.1 Caractérisation du milieu urbain . . . 34
2.1.2 Caractérisation du milieu forestier . . . 35
2.1.3 Cartographie des éléments anthropiques en contexte agricole . 36
2.1.4 Cartographie en milieu côtier . . . 38
2.2 Besoins en production de Modèles Numériques de Terrain (MNT) fins . . . 38
2.3 Les données de télédétection utilisées . . . 39
2.3.1 Les images optiques . . . 39
2.3.2 Les données LiDAR . . . 41
2.3.3 Complémentarité des données optiques et LiDAR . . . 43
2.4 Besoins Vs. Données . . . 44
2.5 Les verrous scientifiques . . . 46

3 Bilan des activités de recherche 49
3.1 Méthodologies d'élaboration de produits cartographiques par télédétection . . . 51
3.1.1 Méthodologie globale . . . 51
3.1.2 Méthodologie de fusion de données multi-sources . . . 53
3.2 Calcul et sélection d'attributs . . . 55
3.2.1 Calcul d'attributs optiques . . . 56
3.2.2 Calcul d'attributs LiDAR . . . 57
3.2.3 Réduction de dimensionnalité . . . 58
3.2.4 Les critères de sélection d'attributs . . . 60
3.2.5 Les méthodes d'optimisation pour la sélection d'attributs . . . 61
3.2.6 Applications . . . 62
3.3 Segmentation . . . 66
3.3.1 Détection d'objets cartographiques . . . 66
3.3.2 Classification orientée objet . . . 67
3.3.3 Mesures de qualité d'une segmentation . . . 68

3.3.4 Méthodes de segmentation 2D 69
3.3.5 Méthodes de segmentation de données 3D 72
3.3.6 Applications 74
3.4 Classification - Apprentissage automatique 78
3.4.1 Les forêts aléatoires 79
3.4.2 Le concept de marge d'ensemble 79
3.4.3 Optimisation de l'ensemble d'apprentissage 80
3.4.4 Applications 82
Conclusion 89

4 Perspectives de recherche **91**
4.1 Axe apprentissage automatique 94
4.1.1 Optimisation de l'ensemble d'apprentissage 94
4.1.2 Intégration de données temporelles 95
4.1.3 Projets futurs 96
4.2 Axe Fusion de données multi-sources 98
4.2.1 Projets futurs 99
4.3 Axe Segmentation sémantique et méthodes d'évaluation appliquées 101
4.3.1 Projets futurs 102
Conclusion 102

5 Annexes **103**
5.1 Axe Sélection d'attributs 103
5.1.1 Sélection d'attributs pour la modélisation de la structure forestière 103
5.1.2 Sélection d'attributs pour la fusion de données optique et LiDAR 119
5.1.3 Sélection de bandes utiles pour la caractérisation de matériaux urbains 138
5.2 Comparaison de données optiques et LiDAR 146
5.3 Axe apprentissage automatique 157
5.3.1 Classification de dégâts de tempête sur la forêt des Landes . . 175
5.4 Axe segmentation 196

Bibliographie **205**

Liste des figures

2.1 Différentes formes d'onde en fonction de la taille du faisceau laser émis. (a) Capteur aéroporté : chaque objet atteint contribue spécifiquement à la forme d'onde. Un capteur multi-échos se contentera de détecter un faible nombre de distances : le premier écho (bleu foncé), le dernier (bleu clair), voire des échos intermédiaires (rose). (b) Capteur satellite ou large empreinte : la forme d'onde est l'intégration selon l'altitude des objets présents, avec une influence non négligeable du terrain en zones de pente. [Source [Mallet et al., 2016]]. 42

3.1 Chaîne méthodologique globale pour l'élaboration de produits cartographiques par télédétection. 52
3.2 Schéma global de fusion de données multi-sources et multi-résolutions 53
3.3 Exemples d'attributs LiDAR multi-échos et à retour d'onde complète. 59
3.4 Approche conjointe ascendante et descendante pour la détection de bâtiments. 74
3.5 Détection de toits de bâtiments par relaxation probabiliste 75
3.6 Segmentation de parcellaire agricoles, image Geoeye multispectrale. 76
3.7 Segmentation de nuages de points 3D terrestre. 78
3.8 Classification urbaine par forêts aléatoires. Utilisation du concept de marge pour optimiser l'ensemble d'apprentissage dans le cas de données déséquilibrées. Les faibles marges apparaîssent en noir. 83
3.9 Cartes de pérdictions par AL au niveau L4 (12 ème itération)(a)Stratified Random , (b)Mahalanobis-BT, (e)Mean-shift-Entropy, (f)Mean-shift-BT. (g) Vérité terrain L4. 84
3.10 Schéma d'apprentissage actif pour la détection de limites parcellaires à grande échelle. 85
3.11 Détection des erreurs d'étiquetage. 86
3.12 Cartographie binaire des dégâts de la tempête Klaus, à partir d'images formosat2 (8m) bi-dates 87

Avant-propos

Ce manuscrit présente mon parcours professionnel et une synthèse de mes travaux de recherche actuels et passés et ouvre ensuite les perspectives sur les développements futurs que j'envisage d'approfondir.

Mes travaux de recherche portent sur des développements méthodologiques pour le traitement et l'analyse des images et des nuages de points 3D afin de répondre à des applications environnementales.

Je m'intéresse tout particulièrement aux données de télédétection optique à Haute et Très Haute Résolution spatiale (HR, THR) et spectrale (hyperspectral, superspectral) et aux données actives type LiDAR pour la caractérisation des milieux par nuages de points 3D. Ces données seront présentées dans le chapitre 3.

J'ai fait le choix de regrouper les méthodologies développées suivant trois grands axes de recherche qui correspondent à différentes étapes de la chaîne de traitement des données : 1) calcul et sélection d'attributs, 2) les techniques de segmentation, et 3) les techniques de classification pour la cartographie de l'occupation du Sol (ocs). Ces développements méthodologiques répondent à des besoins thématiques exprimés sur différents milieux : l'urbain, la forêt, le littoral et les milieux cultivés. Le choix des thématique s'est fait en fonction des institutions dans lesquelles j'ai travaillé et des partenariats établis.

Dans le chapitre 1, après la présentation de mon parcours, mon curriculum vitae sera détaillé suivant 4 axes ; les activités pédagogiques, scientifiques, le rayonnement scientifique et enfin les responsabilités scientifiques et administratives. Le chapitre 2 présente les besoins principaux en caractérisation des milieu, qui sont ensuite détaillés par type de milieu. Ces besoins vont justifier le choix des données de télédétection adaptées et vont définir un certain nombre de verrous scientifiques à lever.

Les contributions méthodologiques seront présentées au chapitre 3 qui synthétise mes propres développements méthodologiques ainsi que les travaux de doctorants que j'ai co-encadrés. Les trois axes méthodologiques seront détaillés. Pour chaque axe, les verrous sont présentés. Les méthodologies sont d'abord présentées d'un point de vue théorique et ensuite leurs applications sont présentées suivant différents projets de recherche.

Mes perspectives de recherches seront ensuite présentées dans le chapitre 4 illustrant les deux grands axes de recherche pour la cartographie de l'ocs à grande et large échelle sur lesquels je souhaite continuer à travailler : 1) l'apprentissage automatique avec intégration de données temporelles et 2) Fusion de données multi-sources.

Parcours et curriculum vitae

1

Sommaire

1.1 parcours 5
1.2 Curriculum Vitae 8

Après une présentation chronologique de mon parcours et des différents organismes dans lesquels j'ai pu évoluer, mon curriculum vitae sera présenté suivant 4 axes : les activités d'enseignement, les activités scientifiques, les responsabilités (pédagogiques, administratives et collectives) et enfin le rayonnement scientifique.

1.1 PARCOURS

Ma thèse de doctorat en informatique a été effectuée entre 2001-2005 sur la modélisation 3D de bâtiments à partir d'images à très haute résolution spatiale au sein de l'équipe SIP (Systèmes Intelligents de Perception) de l'université Paris 5 et au laboratoire MATIS (Méthodes d'Analyses pour le Traitement d'Images et la Stéréorestitution) de l'IGN [1] Saint Mandé avec un co-financement du CNES [2]. Le MATIS mène des recherches en mathématiques et en informatique appliquées à la photogrammétrie, la vision par ordinateur et la télédétection pour l'imagerie multi-capteurs (optique, LiDAR, radar, etc.) et multi-sources (terrestre, aéroportée et/ou satellitaire).

Après deux années (2004-2006) passées en tant que ATER en Informatique à l'université Paris 5, j'ai été recrutée en 2006 en tant que maître de conférences en informatique à l'institut EGID [3]-université bordeaux Montaigne. J'ai été rattachée à l'EA 4134 GHYMAC [4] présentant deux axes : géologie sédimentaire, hydrologie et matériaux. Cette équipe étant très orientée vers les géosciences avec très peu de collègues informaticiens CNU 27, j'ai toujours maintenu mes collaborations avec l'IGN au travers d'une convention de chercheur invité afin de continuer mes développements méthodologiques en traitement d'images pour la télédétection. Avec le laboratoire MATIS, j'ai essentiellement travaillé sur le milieu urbain en utilisant des données multi-sources optique et lidar aéroporté et terrestre pour la classification de scènes urbaines. Dans ce cadre, j'ai co-encadré de deux stages de fin d'étude, et deux thèses : 1)la thèse de Li Guo montrant l'apport des données multi-sources optique et LiDAR pour la classification de scènes urbaines via des classifieurs d'ensemble et 2) la thèse d'Arnaud Le-bris sur la sélection de bandes hyperspectrales utiles pour la caractérisation des matériaux urbains. Ces activités d'encadrement se retrouvent dans les trois grands axes méthodologiques du chapitre 3.

A partir de 2008, j'ai souhaité développer des collaborations en Aquitaine pour me rapprocher des thématiques environnementales de l'EA GHYMAC et participer à une dynamique régionale autour de la télédétection. J'ai alors mis en place des collaborations avec l'INRA Bordeaux [5], UMR ISPA (anciennement EPHYSE). Nous avons développé des activités sur la caractérisation du milieu forestier et la détection de changements suite à la tempête Klaus de 2009 sur la forêt des Landes. Nous avons eu des collaborations fructueuses, soutenues par le CNES à travers la base de données satellitaires LITTORAL Kalidéos et le financement du projet TOSCA STELLA-THR (réseau Scientifique de Télédétection Aquitain-THR). Ces activités ont conduit au co-encadrement d'un master, d'un CDD et de la thèse de Benoit Beguet sur la quantification et caractérisation de structures forestières à partir d'images à très haute résolution spatiale. La thématique forestière étant prioritaire en Aquitaine pour la filière du bois, nous avons bénéficié du soutien de la région Aquitaine pour la conduite de ces travaux de recherche dans le cadre du projet région STELLA.

Cette thématique forestière a également été renforcée dans mes applications lors de la fusion de l'IGN avec l'IFN [6] en 2012 où j'ai pu développer avec

1. IGN : Institut National de l'information géographique et forestière
2. CNES :Centre National des études Spatiales
3. EGID : Environnement, Géo-ingénierie et Développement
4. GHYMAC : Géosciences, Hydrosciences et Matériaux,Construction
5. INRA : Institut National de Recherche Agronomique
6. IFN :Inventaire Forestier National

le laboratoire MATIS des méthodologies de détection de desserte forestière à partir de données LiDAR par exemple.

En 2011, au niveau Aquitain, nous (consortium de chercheurs en télédétection en Aquitaine) avons créé une AST (Action Scientifique Transverse) Télédétection au sein de l'OASU[7], regroupant les différentes équipes travaillant en télédétection en Aquitaine et dans laquelle j'ai été membre active. Cette AST a été coordonnée par Jean Pierre Wigneron (UMR ISP- INRA Bordeaux) et Philippe Paillou (LAB Laboratoire d'Astrophysique de Bordeaux, Université de Bordeaux). A travers cette AST, le site Aquitain autour du bassin d'Arcachon a été identifié comme un site pilote pour le CNES afin de mutualiser des acquisitions satellitaires sur un même site, entre différentes équipes pour servir trois thématiques phares pour la région Aquitaine (vigne, forêt et littoral). Par conséquence, la disponibilité de la base de données KALIDEOS LITTORAL m'a aussi permise de me rapprocher des collègues de l'UMR CNRS 5805 EPOC (Environnements et Paléoenvironnements Océaniques et Continentaux) et de leur cellule de transfert "Geotransfert" travaillant sur la thématique du littoral. Sur cette nouvelle thématique, j'ai également collaboré avec le laboratoire MATIS sur la classification de nuage de points 3D LiDAR sur le littoral pour la classification des points eau/non et l'extraction du trait de côte. D'un point de vue enseignement, cette disponibilité des données m'a permise également de rajouter la thématique littoral à la formation en télédétection de nos étudiants. Depuis deux ans, j'ai mis en place un module intégré sur les plans de prévention des risques littoraux faisant intervenir différents acteurs tel que la DREAL[8], le BRGM[9] ou encore des industriels autour de cette thématique.

Cette AST-Télédétection nous a permis une meilleure coordination entre équipes bordelaises et une plus grande visibilité au niveau national. Nous avons réussi par conséquence à lancer deux projets communs autour de STELLA (réseau Scientifique de TéLédétection Aquitain) ; le projet STELLA de la région Aquitaine et le projet TOSCA/CNES STELLA-THR utilisant des données multi-sources (optique/radar) à très haute résolution pour la caractérisation et le suivi des milieux.

En 2011-2012, nous avons connu de grands changements au sein de notre structure d'enseignement et de recherche qui ont affecté mes orientations sur les deux plans. Le premier changement primordial au niveau de notre offre de formation a eu lieu en 2011 quand l'institut EGID a été habilité par la Commission des Titres d'Ingénieurs et est donc devenu l'école d'ingénieur ENSEGID[10], rattachée à l'institut Polytechnique de Bordeaux (IPB ensuite Bordeaux INP). Ceci a constitué un changement important au niveau de l'orientation de nos enseignements, où nous avons décidé, en fonction des métiers et des débouchés en environnement, de nous focaliser sur la télédétection et les sciences de l'information géographique et d'appuyer plus les applications thématiques que les aspects théoriques en traitement d'images et du signal. En 2012, nous avons également connu un recentrage de nos activités de recherche sur les

7. OASU : Observatoire Aquitain de des Sciences de l'Univers
8. DREAL : Direction Régionale de l'Environnement, de l'Aménagement et du Logement
9. BRGM : Bureau de Recherches Géologiques et Minières
10. ENSEGID : Ecole Nationale Supérieure en Environnement, Géoressources et Ingénierie du Développement durable

géosciences au sein de la nouvelle équipe d'accueil EA 4592 Géoressources et Environnement G&E. Les activités de notre équipe portent désormais sur deux axes : 1) systèmes sédimentaires réservoirs et 2) transferts et interactions dans les systèmes aquifères.

A ce moment là, j'ai ressenti le besoin de faire une conversion thématique afin de me rapprocher des thématiques de l'EA G&E et surtout de l'axe 2 autour des ressources en eau. Suite à mon congé de maternité, en 2012, j'ai bénéficié d'une mise en délégation de deux ans 2012-2014 à l'UMR 144 LISAH [11] (SupAgro [12], INRA, IRD [13]) avec une affectation géographique à l'IRD Tunis. Les recherches du LISAH portent sur l'hydrologie, le transfert de contaminants et l'érosion des sols dans les paysages cultivés fortement aménagés, et sur la caractérisation des propriétés spatiales et l'évolution de ces paysages. C'est sur ce dernier axe que je me suis positionnée.

Ces deux années passées en tant que chargé de recherche à temps plein, dans une unité pluridisciplinaire m'ont amenée à me focaliser sur la thématique des ressources en eau et en sol, enjeu primordial dans les bassins méditerranéens. Je me suis alors concentrée sur la caractérisation des paysages en milieu cultivé et de leurs éléments constitutifs (parcellaire, occupation du sol, états de surface, ..). Dans ce cadre, j'ai assuré l'encadrement de sept stagiaires sur la classification de l'occupation du sol, des états de surface ou encore sur la délimitation du parcellaire agricole. Sur cette thématique, j'ai également proposé le sujet de thèse d'Ines Ben Slimène (2014-2017) pour répondre à un verrou scientifique qui est le problème de passage à l'échelle de classifieurs supervisés en utilisant des techniques d'apprentissage actif (*cf.* Section 3.4.3).

Depuis 2014, j'ai repris mes activités d'enseignement à l'ENSEGID et de recherche dans l'EA G&E. En 2015-2016, j'ai identifié des apports spécifiques en géosciences grâce aux images par drone et la flexibilité des capteurs embarqués. Ainsi, j'ai essayé de promouvoir l'utilisation de l'imagerie par drone pour différentes applications à travers la co-organisation du colloque "drones et moyens légers aéroportés et leurs applications géospatiales en recherche en Tunisie en Novembre 2015 ou encore l'organisation du 1er workshop de l'EA G&E en Mars 2016 sur l'apport de l'imagerie par drone pour les sciences de la terre, workshop impliquant les chercheurs aquitains en télédétection, géosciences mais surtout les industriels. Plusieurs idées de projets collaboratifs ont émergé et seront développés à court/moyen terme. A titre d'exemples, quelques points de convergence entre la télédétection à très haute résolution spatiale et spectrale et les géosciences peuvent être l'utilisation de l'imagerie par drone pour la modélisation 3D d'affleurements (techniques de photogrammétrie), ou encore la cartographie de corps sédimentaires en milieu littoral (par imagerie hyperspectrale), ou encore l'utilisation de l'imagerie thermique pour cartographier les échanges aquifères-rivières. Étant données les institutions dans lesquelles j'ai travaillé (MATIS (IGN), EA G&E (Bordeaux INP-Université Bordeaux Montaigne), UMR LISAH (IRD/INRA/SupAgro), j'ai toujours côtoyé divers thématiciens et je me suis tout de suite positionnée, en recherche appliquée, sur des développements méthodologiques en traitement de données de télédétection pour répondre à des enjeux environnementaux et sociétaux et surtout apporter des réponses à des besoins

11. LISAH : Laboratoire d'étude des Interactions Sol-Agrosystèmes-Hydrosystèmes
12. SupAgro : Institut National d'études supérieures agronomiques de Montpellier
13. IRD : Institut de recherche pour le Développement

thématiques. Étant en section informatique, mon positionnement en recherche appliquée entre thématiciens et informaticiens plus fondamentaux, constitue une forte spécificité de mon profil.

Dans ce qui suit, sera exposé mon curriculum vitae mettant en avant mes activités pédagogiques, scientifiques, mes différentes responsabilités aussi bien pédagogiques, collectives que scientifiques et enfin le rayonnement scientifique.

1.2 Curriculum Vitae

CURRICULUM VITAE- NESRINE CHEHATA

IDENTIFICATION

Mme Nesrine CHEHATA, Née le 07/11/1978, Mariée, 1 enfant
Maître de conférences en Informatique, 5ème échelon, CNU 27
Spécialité: Traitement d'images pour la télédétection
Etablissement: Bordeaux INP - ENSEGID
Unité de recherche: EA 4592 Géoressources et Environnement (G&E)- Bordeaux INP/ Université Bordeaux Montaigne
2012-2014: Mise en délégation à l'UMR LISAH 144 (IRD/INRA/SupAgro) - Affectation à l'IRD Tunis
Depuis 2006: Chercheur invité au laboratoire MATIS - IGN Saint Mandé

FORMATION

Habilitation à diriger des recherches en Informatique (A soutenir) *20 Mars2017*
Contributions méthodologiques pour la caractérisation des milieux par télédétection optique THR et LiDAR
Composition du jury:
Jocelyn Chanussot (INP Grenoble - section 27) : Rapporteur
Laurent Wendling (Université Paris 5 - section 27): Rapporteur
Florence Tupin (Université Paris 5 - section 61): Rapporteur
Pierre Gançarski(Université de Strasbourg - section 27): Examinateur
Christian Heipke (Professeur à Leibnitz Universitat): Examinateur
Nicolas Paraoditis (Directeur de recherche à l'IGN): Examinateur
Qualification aux sections CNU 27 et 61 *2005*
CNU 27 Informatique
CNU 61 Génie informatique, automatique et traitement du signal
Doctorat en Informatique. *2001-2005*
Modélisation 3D de scènes urbaines à partir d'images satellitaires à très haute résolution
Mention Très honorable
Bourse (CNES/IGN), Université Paris Descartes, Paris 5,
Présidente du jury: Nicole Vincent (Université Paris 5 - section 27)
Directeur de thèse: Georges Stamon (Université Paris 5 - section 27)
Rapporteur : Mathieu Cord (ENSEA - section 27)
Rapporteur: Henri Maître (Télécom Paris - section 61)
Examinateur: Marc Pierrot-Deseilligny (MATIS/IGN)
Examinateur: Franck Jung (MATIS/IGN)
Examinateur: Gilbert Pauc (CNES)
Diplôme d'Études Approfondies en Photonique, Image et Cybernétique Option « Image ». *2000-2001*
Mention Bien
Université Louis Pasteur, Strasbourg I
Diplôme d'ingénieur Télécom Physique Strasbourg, spécialité traitement d'image *1998-2001*
Mention Bien
Anciennement Ecole Nationale Supérieure de Physique de Strasbourg
Classes préparatoires MP(Math-Physique) *1996-1998*
Lycée du Parc, Lyon, France
Bourse d'étude de l'état Tunisien
Diplôme de baccalauréat Section Math (S). Lauréate du Lycée Menzah 6 - Tunis *1996*
mention Très Bien

EXPÉRIENCE

Maître de conférences en informatique, 5 ème échelon 2014 - maintenant
ENSEGID - Bordeaux INP

· Ecole Nationale Supérieure en Environnement, Géoressources et ingénierie du Développement durable

Délégation en tant que chargé de recherches CR1, section Informatique 2012 - 2014
UMR LISAH 144 (SupAgro, INRA, IRD) avec affectation à l'IRD Tunis *Institut de Recherche pour le Développement*

Maître de conférences en informatique 2011 - 2012
ENSEGID - IPB (Institut Polytechnique de Bordeaux)

Maître de conférences en informatique 2006 - 2011
Institut EGID (EGID Environnement, Géo- ingénierie et Développement)

Université Bordeaux 3

Attaché Temporaire d'Enseignement et de Recherche (ATER) en informatique 2004-2006
A mi-temps, université Paris Descartes - Paris 5

ACTIVITÉS PÉDAGOGIQUES

COURS MAGISTRAUX ET TP

Maître de conférences en informatique ENSEGID-Bordeaux INP 2011, 2014 - maintenant

· 192 h éq TD par an
· Traitement d'images, Télédétection, géomatique
· Logiciels: Matlab, ENVI, ArcGis,QGIS

Maître de conférences en informatique EGID 2006 - 2011

· 192 h éq TD par an
· programmation C/C++, statistiques et analyse de données, Traitement d'images, Télédétection,
· Logiciels: Matlab, excel, ImageJ, ENVI

Demi-ATER en informatique à l'université Paris 5 2004-2006

· 192 h éq TD
· Analyse d'images, Programmation orientée objet, Programmation Turbo Pascal

Vacataire en Informatique 2001-2004
Université Paris 5, ESGT (Ecole Supérieure des Géomètres et Topographes),
ECE Paris (Ecole Centrale d'Electronique) *180h*

· Télédétection, Programmation Turbo Pascal, C++, théorie des graphes

Vacataire en informatique 2002
Faculté de médecine de Sousse - Tunisie *40h*

· Traitement des images médicales pour améliorer leur interprétation , logiciel: imageJ
· public 15 médecins pratiquants

ANIMATIONS DE FORMATIONS

Formation organisée par le CNCT et le CRETEAN 2014
Tunis, Tunisie *15h*

- Animation de la formation "Technical training course on hyperspectral remote sensing imagery analysis - advanced level" organisée par le CNCT(Centre National de Cartographie et Télédétection et le CRETEAN (Centre régional de télédétection des états d'Afrique du Nord)
- Niveau avancé : prétraitement des données hyperspectrales, classification, démixage

Formation FOSTIT, réseau Télédétection AUF 2013
Tunis, Tunisie *20h*

- Co-animation de la "Formation sur les Outils et Systèmes de Traitement d'Images de Télédétection", Tunisie - FORSTIT 2013
- Niveau avancé: Traitement d'images, télédétection, OrfeoToolBox, Monteverdi

COURS-CONFÉRENCE

2 ème SummerSchool Bordeaux INP sur le développement durable en Aquitaine 2016

- *"Contribution of multisource remote sensing for forest sustainable development"*
- Accueil d'étudiants de l'université RMIT Melbourne, Bordeaux.

ENIT (Ecole Nationale des ingénieurs de Tunis), Tunisie 2016

- *"Apport des données Lidar en génie hydraulique"*
- *"Apport des données Lidar en génie civil"*
- 2 eme année Génie hydraulique, 2 ème année génie civil

1 ère SummerSchool Bordeaux INP sur le développement durable en Aquitaine 2015

- *"Contribution of remote sensing for forest sustainable development in Aquitaine Region"*
- Accueil d'étudiants de l'université RMIT Melbourne, Bordeaux.

Ecole d'hiver "Techniques d'observation et méthodes d'analyse pour la gestion de l'eau dans les bassins versants agricoles méditérannéens" 2015

- *"Chaîne de traitement pour la classification de l'occupation du sol à grande échelle"*

SUP'COM (Ecole Supérieure des Communications de Tunis), Tunisie 2009,2012,2013,2014,2015

- *"Lidar et applications environnementales"*
- master – option contenu multimédia

ENIT (Ecole Nationale des Ingénieurs de Tunis), Tunisie 2007,2008

- *"Imagerie satellitaire et applications"*
- master ATS : Automatique et Traitement du Signal.

INRA Grignon 2002

- *"Reconstruction 3D des bâtiments en milieu urbain"*
- mastère SILAT (Systèmes d'Informations Localisées pour L'Aménagement des Territoires)

ACTIVITES SCIENTIFIQUES

Mes activités de recherche portent sur des développements méthodologiques en traitement et analyse des données de télédétection pour la caractérisation et le suivi des milieux. Je me spécialise dans le traitement des images optiques à Très Haute Résolution spatiale ou spectrale et des nuages de points 3D Lidar.
Mes travaux peuvent être regroupés suivant 3 axes de recherche méthodologiques:

1. les méthodes de sélection d'attributs
2. les méthodes de segmentation 2D/3D pour la détection d'objets cartographiques et la classification
3. les méthodes d'apprentissage par ensembles de classifieurs pour la cartographie de l'occupation du sol

   Au niveau des thématiques, j'ai essentiellement travaillé sur le milieu urbain et littoral en collaboration avec l'IGN et le CNES, le milieu forestier en collaboration avec l'INRA Bordeaux pour la caractérisation et la détection de changements des structures forestières et enfin sur les milieux cultivés en collaboration avec l'UMR LISAH (IRD Tunis) et l'INAT (Institut National Agronomique de Tunisie).

Mots-clé: Traitement d'images, télédétection, classification, sélection d'attributs, classifieurs d'ensemble, segmentation, LiDAR, imagerie optique THR, hyperspectral
Applications: urbain, forêt, agriculture

RATTACHEMENT SCIENTIFIQUE

Enseignant-chercheur 2014-maintenant
EA 4592 Géoressources et Environnement, Université Bordeaux Montaigne

· chercheur invité au laboratoire MATIS (Méthodes d'Analyses pour le Traitement d'Images et la Stéréorestitution), IGN, Action de Recherche RISOTO

Chargé de recherche 2012-2014
UMR LISAH 144 (IRD, INRA, SupAgro)

· Délégation, affectation géographique à l'IRD Tunis

Enseignant-chercheur 2006-2012
EA 4592 Géoressources et Environnement, Université Bordeaux Montaigne

· chercheur invité au laboratoire MATIS, IGN

ATER - Chercheur associé 2004-2006
Laboratoire MATIS, IGN

Doctorante 2001-2005
Laboratoire MATIS, IGN

AXES DE RECHERCHE ET MÉTHODOLOGIES

Axe 1: Sélection d'attributs

· Stratégies séquentielles
· Stratégies par ensembles de classifieurs
· Stratégies stochastiques
· Applications: caractérisation de variables biophysiques forestières , classification multi-sources optique et LiDAR, extraction de bandes utiles en imagerie hyperspectrale pour l'optimisation de la configuration d'un capteur

Axe 2: Les méthodes de segmentation 2D/3D

- segmentation multi-échelles
- segmentation par coupes de graphe
- segmentation par relaxation probabiliste
- Applications: segmentation de nuages de points 3D, détection de bâtiments, délimitation du parcellaire agricole, détection de routes à partir de données LiDAR

Axe 3: Méthodes d'apprentissage

- ensemble de classifieurs, marge d'ensemble
- apprentissage actif
- optimisation de l'apprentissage
- partitionnement hiérarchique
- Applications: production de MNT LiDAR, délimitation du parcellaire agricole, détection de changement, classification de l'occupation du sol à large échelle

PUBLICATIONS ET PRODUCTION SCIENTIFIQUE

74 items : 4 chapitres de livres, 9 articles de revues internationales, 2 soumis, 7 articles de revues nationales françaises, 48 communications internationales avec actes (dont 17 de rang B, 12 de rang C), et 3 communications nationales françaises avec actes Tous les facteurs d'impact IF sont donnés pour l'année 2015.
H-index : 6 - 237 citations (Scopus)
H-index: 11 , Hi10-index: 13 - 530 citations (Google scholar)

PUBLICATIONS MAJEURES

[Guo et al, 2011] [6] Cet article est le 5 ème article le plus cité de la revue IJPRS (IF 2015: 4.188). Article de référence pour l'utilisation des Random Forests (RF) pour la sélection d'attributs en télédétection multi-sources (lidar et multispectral) pour la classification urbaine. L'originalité méthodologique réside dans l'utilisation de la marge d'ensemble comme critère de pertinence des attributs et l'apport thématique a été de montrer l'intérêt des données lidar à retour d'onde complète (*Full WaveForm*) en urbain.
Contribution: Application thématique thèse Li GUO, rédaction de l'article, analyse des résultats

[Chehata et al, 2014][4] présente une approche région quasi non supervisée de détection de changements sur des images satellite. Une sélection automatique d'attributs optimise la segmentation et la classification, de type mean shift. Un nouveau descripteur spatio-temporel, basé sur le taux de fragmentation des régions détectées, permet de réaliser une classification binaire des changements.
Contribution: encadrement principal du stage de Camille Orny, Rédaction de l'article

[Beguet et al.,2014][3] présente l'apport de l'imagerie THR pléiades pour la caractérisation de variables forestières par régressions linéaires multiples. Un critère innovant de sélection d'attributs a été proposé et on a montré l'intérêt des attributs multi-échelles et multi-résolutions pour la modélisation de variables biophysiques. La robustesse de la méthode a été validée sur plusieurs dates intra-saison.
Contribution: Encadrement de la thèse de Benoit Beguet. Commande des données multi-temporelles et validation de la robustesse de la méthode.

[Le-Bris et al.,2015] [5] porte sur l'axe de recherche sélection d'attributs et présente une méthode innovante de partionnement (clustering) hiérarchique de bandes hyperspectrales afin d'optimiser le choix des bandes utiles en prenant en compte leur position dans le spectre mais également leur largeur. L'optimisation est faite par algorithmes génétiques. L'application finale consiste à optimiser la configuration d'une caméra superspectrale (20 bandes), qui sera produite par l'IGN/ laboratoire LOEMI. Un article de revue internationale est en cours de rédaction sur ce travail.

[Ferraz et al.,2016][5] choisi comme "featured article" du journal IJPRS (publié en libre accès pendant 1 an). L'article présente une méthode automatique de détection, à large échelle, des routes en milieu forestier et montagneux par données LiDAR.
Contribution: Utilisation des classifieurs d'ensemble (Random Forests) sur de grands volumes de données, utilisation de la marge d'ensemble et optimisation par coupe de graphes, évaluation des résultats.

[Smeeckaert et al.,2014] [5] présente une méthode automatique pour la classification, à large échelle des zones côtières par des données LiDAR.
Contribution: à la mise en place de la méthodologie, la définition des descripteurs LiDAR appropriés(basés sur la densité de points) et à la partie régularisation par relaxation probabiliste des cartes de classification finales.

LISTE DES PUBLICATIONS

Chapitres de livres

[1] C. Mallet, **N. Chehata**, and J-S. Bailly. *Airborne LiDAR data processing*. Remote Sensing of Land Surfaces, Volume 1: Remote sensing techniques, principles and methods, chapter 06 ISTE Editor, Elsevier, Land- Environment System, 2016

[2] X.Briottet, **N.Chehata**, R. Oltra-Carrio, A.Le Bris, and C. Weber. *Optical remote sensing in urban environements*. Remote Sensing of Land Surfaces, Volume 4: Land Surface Remote Sensing in Urban and Coastal Areas, chapter 01 ISTE Editor, Elsevier, Land- Environment System, 2016

[3] C. Mallet, **N. Chehata**, and J-S. Bailly. *Méthodes de traitement de données LIDAR*. Télédétection pour l'Observation des Surfaces Continentales, Tome 1: Techniques de télédétection : principes et méthodes, chapitre 06 ISTE Editeur, Elsevier, Système Terre – Environnement, 2016

[4] X.Briottet, **N.Chehata**, R. Oltra-Carrio, A.Le Bris, and C. Weber. *Application de l'imagerie optique aux milieux urbains*. Télédétection pour l'Observation des Surfaces Continentales, Tome 4: Applications thématiques: Environnement et Risques ISTE Editeur, Elsevier, Système Terre – Environnement, 2016

Articles de journaux internationaux à comité de lecture

[5] I. Ben Slimene-Ben Amor, **N. Chehata**, P. Lagacherie, J-S. Bailly , and Riadh Imed Farah. Parcel-based active learning for large extent cultivated area mapping *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (IF=2.145)*, 2017. **soumis**.

[6] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Which feature selection score to design a superspectral sensor for classification ? *{ISPRS} Journal of Photogrammetry and Remote Sensing* (IF=4.188), 2017. **soumis**.

[7] G. Sofia, J.-S. Bailly, **N. Chehata**, P. Tarolli, and F. Levavasseur. Comparison of pleiades and lidar digital elevation models for terraces detection in farmlands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (IF=2.145)*, pp(99):1–10, 2016.

[8] A. Ferraz, C. Mallet, and **N. Chehata**. Large-scale road detection in forested mountainous areas using airborne topographic lidar data. *{ISPRS} Journal of Photogrammetry and Remote Sensing (IF=4.188)*, 112:23–36, 2016.

[9] C. Mallet, **N. Chehata**, and G. Mercier Theme Issue "Multitemporal remote sensing data analysis. *{ISPRS} Journal of Photogrammetry and Remote Sensing* (IF=4.188), 107:1–2, 2015

[10] B. Beguet, D. Guyon, S. Boukir, and **N. Chehata**. Automated retrieval of forest structure variables based on multi-scale texture analysis of {VHR} satellite imagery. *{ISPRS} Journal of Photogrammetry and Remote Sensing (IF=4.188)*, 96(0):164–178, 2014.

[11] **N. Chehata**, C. Orny, S.Boukir, and D. Guyon. Object-based change detection in wind-storm damaged forest using high resolution multispectral images. *International Journal of Remote Sensing (IF=1.652)*, 35(13):4758–4777, 2014.

[12] J. Smeeckaert, C. Mallet, N. David, **N. Chehata**, and A. Ferraz. Large-scale classification of water areas using airborne topographic lidar data. *Remote Sensing of Environment (IF=5.881)*, 138:134 – 148, 2013.

[13] L. Guo, **N. Chehata**, C. Mallet, and S. Boukir. Relevance of airborne lidar and image data for urban scene classification using random forests. *{ISPRS} Journal of Photogrammetry and Remote Sensing (IF=4.188)*, 66(1):56–66, Janvier 2011.

[14] F. Bretar and **N. Chehata**. Terrain modelling from lidar range data in natural landscapes: a predictive and bayesian framework. *IEEE Transactions on Geoscience and Remote Sensing (IF=3.514)*, 48(3):1568–1578, 2010.

[15] **N. Chehata**, F. Jung, and G. Stamon. A graph cut optimization guided by 3D-features for surface height recovery. *ISPRS Journal of Photogrammetry and Remote Sensing (IF=4.188)*, 64(2):193–203, 2009.

Articles de journaux nationaux avec comité de lecture

[16] **N. Chehata**, K. Ghariani, A. Le Bris, and P. Lagacherie. Délimitation des parcelles agricoles par classification d'images pléiades. *Revue Française de Photogrammétrie et de Télédétection*, 209(165-171), 2015.

[17] A. Ferraz., C. Mallet, and **N. Chehata**. Détection de la desserte forestière en milieu montagneux par lidar aéroporté. *Revue Française de Photogrammétrie et de Télédétection*, 211-212,pp. 103-117, 2015.

[18] B. Beguet., **N. Chehata**, D. Guyon and S. Boukir. Quantification et cartographie de la structure forestière à partir de la texture des images pléiades *Revue Française de Photogrammétrie et de Télédétection*, 208(83-88), 2014.

[19] S. Boukir, C. Orny, **N. Chehata**, and D. Guyon D. Détection de changements structurels sur des images satellite haute résolution. application en milieu forestier. *Revue Traitement du signal*, volume 30(6/2013):401–428, 2013.

[20] F. Bretar and **N. Chehata**. Génération de modèles numériques de terrain par fusion de données lidar et image. *Revue Traitement du Signal*, 26(2), 2009.

[21] **N. Chehata** and G. Stamon. Évaluation de produits cartographiques: un système opérationnel à différents niveaux de contrôle. *Revue Traitement du Signal*, 24(5):371–382, 2007.

[22] **N. Chehata**, M. Pierrot-Deseilligny, and G. Stamon. Génération de mne hybrides (raster/vecteur) : optimisation à base de flots de graphes contrainte par des primitives. *Revue Française de Photogrammétrie et de Télédétection*, 176 (2004-4):30–41, 2004.

Actes de conférences internationales avec comité de lecture

[23] W. Ouerghemmi, A. Le-Bris, **N. Chehata**, and C. Mallet A two-step decision fusion strategy: application to hyperspectral and multispectral images for urban classification, ISPRS Workshop In *CMRT City Models, Roads and Traffic*, Jun 2017, Hannover, Germany

[24] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Hierarchically exploring the width of spectral bands for urban material classification In *JURSE Joint Urban Remote Sensing Event (Rank B)*, Mars 2017, Dubai, UAE

[25] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Spectral band selection for urban material classification using hyperspectral libraries In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (Rank B)*, volume III-7, XXIII ISPRS Congress, pages 33–40, Jul 2016, Prague, Czech Republic

[26] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Extraction of optimal spectral bands using hierarchical band merging out of hyperspectral data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (Rank B)*, Volume XL-3/W3, pages 459–465, Oct 2015.

[27] I. Ben Slimene-Ben Amor, **N. Chehata**, P. Lagacherie, J-S. Bailly , and Riadh Imed Farah. Can we automatically choose best uncertainty heuristics for large margin active learning? In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2015 (Rank C)*, 2015.

[28] C. Cazals, H. Benelcadi, P.-L. Frison, G. Mercier, C. Lardeux **N. Chehata**, Champion and J.-P. Rudant. Contribution of textural information from terrasar-x image for forest mapping. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2015 (Rank C)*, 2015.

[29] A. Le Bris, **N. Chehata**,X. Briottet, and N. Paparoditis. A random forest class memberships based wrapper band selection criterion : Application to hyperspectral. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2015 (Rank C)*, 2015.

[30] B. Beguet, D. Guyon, S. Boukir, and **N. Chehata**. Classification of forest structure using very high resolution pleiades image texture. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2014 (Rank C)*, pages 2324–2327, 2014.

[31] I. Ben Slimene, **N. Chehata**, I.R. Farah, and P. Lagacherie. Uncertainty heuristics of large margin active learning for hyperspectral image classification. In *IEEE IPAS'14 International Image Processing applications and systems conference 2014 (Rank C)*, pages 1–6, 2014.

[32] A. Ferraz, C. Mallet, and **N. Chehata**. Large-scale road network extraction in forested mountainous areas using airborne laser scanning data. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2014 (Rank C)*, pages 4315–4318, 2014.

[33] K. Ghariani, **N. Chehata**, A. Le Bris, and P. Lagacherie. Agricultural field delimiitation using active learning and random forests margin. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2014 (Rank C)*, pages 1717–1720, 2014.

[34] M.M. Sidi Youssef, C. Mallet, **N. Chehata**, A. Le Bris, and A. Gressin. Combining top-down and bottom-up approaches for building detection in a single very high resolution satellite image. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2014 (Rank C)*, pages 4820–4823, 2014.

[35] **N. Chehata**, A. Le Bris, and S. Najjar. Contribution of band selection and fusion for hyperspectral classification. In *IEEE WHISPERS*, Lausanne, Suisse, Jun 2014.

[36] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Identify important spectrum bands for classification using importances of wrapper selection applied to hyperspectral data. In *IEEE*

IWCIM (International Workshop on Computational Intelligence for Multimedia Understanding), 2014.

[37] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Use intermediate results of wrapper band selection methods: A first step toward the optimization of spectral configuration for land cover classifications. In *IEEE WHISPERS*, Lausanne, Suisse, Jun 2014.

[38] B. Beguet, Boukir S, Guyon D, and **N. Chehata**. Modelling-based feature selection for classification of forest structure using very high resolution multispectral imagery. In *IEEE SMC 2013 (IEEE International Conference on Systems, Man, and Cybernetics) (Rank B)*, pages 4294–4299, Manchester, UK, October 2013 2013.

[39] S. Boukir, L. Guo, and **N. Chehata**. Classification of remote sensing data using margin-based ensemble methods. In *IEEE International Conference on Image Processing ICIP 2013 (Rank B)*, pages 2602–2606, Melbourne, Australia, Sep 2013 .

[40] **N. Chehata**, A. Le Bris, and P. Lagacherie. Comparison of VHR panchromatic texture features for tillage mapping. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2013 (Rank C)*, pages 3128–3131, 2013.

[41] A. Le Bris, F. Tassin, and **N. Chehata**. Contribution of texture and red-edge band for vegetated areas detection and identification. In *IEEE International Geoscience and Remote Sensing Symposium IGARSS 2013 (Rank C)*, pages 4102–4105, 2013.

[42] J. Smeeckaert, C. Mallet, N. David, A. Ferraz, and **N. Chehata**. Large scale classification of coastal areas using ariborne topographic lidar data. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2013 (Rank C)*, pages 61–64, 2013.

[43] B. Beguet, **N. Chehata**, S. Boukir, and D. Guyon. Retrieving forest structure variables from very high resolution satellite images using an automatic method. In *XXII International Society for Photogrammetry & Remote Sensing Congress, ISPRS Annals of Photogrammetry, Remote Sensing and the Spatial Information Sciences*, volume I-7, pages 1–6, Melbourne, 2012.

[44] **N. Chehata**, C. Orny, S. Boukir, and D. Guyon. Object-based forest change detection using high resolution satellite images. In *ISPRS PIA 2011 (Photogrammetric Image Analysis)*, volume 38 Part 3/W22, pages 49–54, Munich, 5-7 Octobre 2011.

[45] A. Le Bris and **N. Chehata**. Change detection in a topographic building database using submetric satellite images. In *ISPRS PIA 2011 (Photogrammetric Image Analysis)*, Munich, 5-7 Octobre 2011.

[46] L. Guo, S. Boukir, and **N. Chehata**. Support vectors selection for supervised learning using an ensemble approach. In *IAPR International Conference on Pattern Recognition (ICPR)(Rank B)*, pages 37–40, Istanbul, 23-26 Août 2010.

[47] L. Guo, **N. Chehata**, and S. Boukir. A two-pass random forests classification of airborne lidar and image data on urban scenes. In *IEEE International Conference on Image Processing (ICIP) (Rank B)*, pages 1369–1372, Hong Kong, 26-29 Septembre 2010 2010.

[48] **N. Chehata**, L. Guo, and C. Mallet. Airborne lidar feature selection for urban classification using random forests. In *Laserscanning 2009, IAPRS*, volume 38 (Part 3/W8), pages 207–212, Paris, sep 2009.

[49] **N. Chehata** and F. Bretar. Génération de modèles numériques de terrain à partir de données lidar par approche multi-résolutions. In *Traitement et Analyse de l'Information: Méthodes et Applications (TAIMA'09)*, Hammamet, 4-9 Mai 2009.

[50] **N. Chehata**, L. Guo, and C. Mallet. Contribution of airborne full-waveform lidar and image data for urban scene classification. In *IEEE International Conference on Image Processing (ICIP) (Rank B)*, Caire, 7-10 Novembre 2009.

[51] **N. Chehata** and F. Bretar. Terrain modeling from lidar data: Hierarchical K-means filtering and markovian regularization. In *IEEE International Conference on Image Processing (ICIP) (Rank B)*, San Diego, 12-15 Octobre 2008.

[52] **N. Chehata**, N. David, and F. Bretar. Lidar data classification using hierarchical K-means clustering. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 37 (Part 3B), pages 325–330, Pekin, Juil 2008.

[53] F. Bretar and **N. Chehata**. Digital terrain model on vegetated areas: Joint use of airborne lidar data and optical images. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W49A), Munich, 19-21 Septembre 2007.

[54] **N. Chehata**, M. Pierrot Deseilligny, and G. Stamon. Hybrid digital elevation model production guided by 3D-primitives: a global optimization algorithm using graph cuts. In *International Conference on Image Processing (Rank B)*, pages 117–120, 2005.

[55] **N. Chehata**, F. Jung, M. Pierrot-Deseilligny, and G. Stamon. Reconstruction de facettes 3D par approche hiérarchique par régions à partir de couples d'images satellite en milieu urbain. In *International Conference on Sciences of Electronic, Technologies of Information and Telecommunications (SETIT)*, Sousse, Tunisia, Mars 2004.

[56] **N. Chehata**, F. Jung, M. Pierrot Deseilligny, and G. Stamon. A region-based matching approach for 3D-roof reconstruction from HR satellite stereo pairs. In *DICTA (Rank B)*, pages 889–898, 2003.

[57] **N. Chehata**, M. Pierrot Deseilligny, F. Jung, and G. Stamon. Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings. In *MVA (Machine Vision Applications) (Rank C)*, pages 636–640, 2002.

Actes de conférences nationales avec comité de lecture

[58] M.M. Sidi Youssef, C. Mallet, **N. Chehata**, A. Le Bris, and A. Gressin. Détection de bâtiments à partir d'une image satellitaire par combinaison d'approches ascendante et descendante. In *RFIA'14 (19 ème congrès national sur la Reconnaissance de Formes et l'Intelligence Artificielle)*, Rouen, Jun 2014.

[59] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Optimisation de configuration spectrale pour la classification de données hyperspectrales. In *3ème colloque scientifique SFPT-GH (Société Française de Phogrammétrie et Télédetection, Groupe Hyperspectral)*, IGESA Porquerolles., May 2014.

[60] S. Boukir, C. Orny, **N. Chehata**, and D. Guyon. Classification multitemporelle des dégâts de tempête en forêt sur des images satellite haute résolution. In *23ème Colloque GRETSI - Traitement du Signal et des Images*, Bordeaux, 5-8 Septembre 2011.

[61] S. Liberge, B. Soheilian, **N. Chehata**, and N. Paparoditis. Extraction of vertical objects in 3D point clouds. In *ISPRS Technical Commission III Symposium on Photogrammetry Computer Vision and Image Analysis*, Paris, 1-3 Septembre 2010.

Articles de conférences internationales sans comité de lecture

[62] I. Ben Slimene, **N. Chehata**, I.R. Farah, P. Lagacherie and J.S Bailly Agricultural Land cover mapping by active learning from multispectral spot-7 satellite image. In *International Conference Exhibition Advanced Geospatial Science Technology (TeanGeo 2016)*,2016

[63] D. Feurer, S. Massuel, A. Sefi, **N. Chehata**and A. Ezzine Which 3D changes can be assessed from SfM processing of historical aerial imagery ? In *International Conference Exhibition Advanced Geospatial Science Technology (TeanGeo 2016)*, 2016.

[64] J.S. Bailly, G.Sofia, **N. Chehata** and P. Tarolli. Farmland terrace slope detection from pleiades digital elevation models. In *EGU 2015-10021, session SSS2.5/GM6.6/HS12.3 Agricultural terraces of the world., eISSN:1607-7962*, 2015

[65] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Spectral band selection for urban material maps using hyperspectral data. In *9th IEEE EARSeL SIG Imaging spectroscopy workshop*, Luxembourg, 2015.

[66] A. Le Bris, **N. Chehata**, X. Briottet, and N. Paparoditis. Very high resolution land cover extraction in urban areas. In *8th IEEE EARSeL Imaging spectroscopy workshop*, Apr 2013.

Rapports techniques

[67] C. Orny, **N. Chehata**, S. Boukir, and D. Guyon. Characterization of maritime pine forest structure changes with VHR satellite imagery : application to the 24th january 2009 windfall damages cartography. Technical report, INRA, CNES, 2010.

[68] **N. Chehata** Hybrid digital elevation model production guided by 3D-primitives: A global optimization algorithm using graph cuts. Technical Report 2006/1, SIP-CRIP5 (Université Paris 5), 2006.

[69] L. Guigues, R. Trias-Sanz, **N. Chehata**, F. Taillandier, and M. Deveau. Segmentation multi-échelles d'images: théorie et applications. *Bulletin d'Information Scientifique et Technique de l'IGN*, 75:41–57, 2005.

[70] **N. Chehata**, F. Jung, and G. Stamon. Interprétation de scènes urbaines à partir d'images satellitaires thr: Reconstruction de facettes 3D et optimisation globale 3D. *Bulletin d'information de l'IGN*, 75:29–40, 2004.

N. Chehata. Extraction de primitives 3D pour la reconstruction 3D de bâti milieu urbain à partir de couples d'images satellites haute résolution. Technical Report 2003/1, SIP-CRIP5 (Université Paris 5), 2003.

Articles de vulgarisation

[71] J-S. Bailly and **N. Chehata**. Comparaison de MNS pléiades-lidar en contexte agricole (languedoc-roussillon). Bulletin d'information, Pôle thématique surfaces continentales,- Théia, N°2, Jul 2014.

[72] Météorites. Un laser dévoile leurs cratères. *Revue SCIENCE&VIE*, 1100:118–119, 2009.

Rapports diplômants

[73] **N. Chehata**. *Modélisation 3D de scènes urbaines à partir d'images satellitaires à très haute résolution*. PhD thesis, Université Paris Descartes, France, Jun 2005.

[74] **N. Chehata**. *Indexation vidéo orientée objet: Combinaison des techniques de segmentation d'images en régions avec l'extraction des descripteurs de couleur.* Rapport de Fin détudes ENSPS et de DEA PIC en photonique, image et cybernétique, spécialité traitement d'images, Université Louis Pasteur, Strasbourg, Sep 2001. 58 p.

ENCADREMENT DOCTORAL ET SCIENTIFIQUE

3 thèses soutenues, 1 thèse en cours, 1 post-doc, 1 CDD, 7 stages de Master 2, 5 stages de fin d'études d'ingénieur et 2 stage de 2 ème année d'école d'ingénieur

ENCADREMENTS DE THESES et POST-DOC

Durée	Etudiant	Sujet	Directeur de thèse	Bourse	% Encadrement	Devenir
2016	Post Doc Walid Ouerghemmi	Fusion décisionnelle de données multi-échelles en milieu urbain	IGN	40%	Post-doc	
2014-2017	Thèse Ines Ben Slimene	Apprentissage actif pour la classification de l'occupation du sol à grande échelle à partir d'images satellitaires à haute résolution spatiale	Riadh Imed Farah (10%) et Philippe Lagacherie(20%)	18 mois de mobilité - Alyssa Erasmus Mundus	70%	-
2011-2015	Thèse Arnaud Le Bris	Occupation du sol à grande échelle, sélection des bandes utiles en imagerie hyperspectrale pour la caractérisation de matériaux urbains	Nicolas Paparoditis (5%) et Xavier Briottet (25%)	IGN	70%	Ingénieur IGN
2010-2014	Thèse Benoit Beguet(*)	Quantification et cartographie de la structure forestière à partir d'images à très haute résolution spatiale	Samia Boukir	CNES/Rég Aquitaine	50 %	ATER ENSEGID puis Ingénieur Téléspazio Aquitaine
2008-2011	Thèse Li Guo	Classifieurs multiples intégrant la marge d'ensemble, Application aux données de télédétection. Margin Framework for ensemble classifiers, Application to remote sensing data	Samia Boukir	Chine	50 %	Post-doc puis CDI Ingénieur informatique -Lille

(*): La thèse de Benoit Beguet (2010-2014) a eu le prix THES'AQT (1 ère édition du forum de l'innovation globale NOVAQT), Aquitaine, France.
Cette thèse a été labellisée par deux pôles de compétitivité: AeroSpace Valley et Xylofutur.

ENCADREMENTS DE CDD ET STAGES

Nature	Durée	Etudiant	Sujet
CDD CNES	Jan-Août 2010	Camille Orny	Caractérisation des couverts forestiers et leur suivi temporel
Master 2	Février-Déc 2017	Nouha Touihri	Classification de types de peuplements forestiers par données multicapteurs (Pléiades et Sentinel)
Master 2	Février-Juin 2016	Meriam Mejri	Mesures de performance de méthodes de segmentation d'images pour l'identification de parcellaires culturaux
Master 2	Juin 2015-Juin 2016	Houcem Daagi	Cartographie de l'occupation du sol et délimitation du parcellaire agricole sur le bassin versant du Lebna , Cap Bon – Tunisie
Master 2	Fev-Juin 2014	Ines Ben Slimene	Apprentissage actif pour la classification de l'occupation du sol à grande échelle
Master 2	Juil 2012-Mai 2013	Nahed Matoussi	Caractérisation des états de surface par imagerie Très haute résolution : cas du BV du Lebna, Cap-Bon Tunisie
Master 2	Avr-Sep 2009	Camille Orny	Détection de changements à partir d'images Formosat pour la cartographie des dégâts de tempête sur la forêt des Landes
Master 2	Fév-Août 2009	Sterenn Liberge	Extraction d'objets verticaux à partir de nuages de points 3D en milieu urbain acquis par laser terrestre embarqué
PFE	Fev-Juin 2017	Simon Bailly	Analyse de séries temporelles d'images Sentinel et intégration de connaissances pour la classification en milieu agricole
PFE	Fev-Juin 2017	Safa Aloui	Détection de strates forestières 3D à partir de données LiDAR
PFE	Fev-Juin 2014	Walid Boucceffa (***)	Mise en place d'une chaîne de classification de l'occupation du sol à grande échelle : application en milieu cultivé, bassin versant du Lebna, Cap Bon
PFE	Jan-Juil 2013	Karim Ghariani	Classifieurs d'ensemble exploitant les faibles marges : application à la détection de limites parcellaires
PFE	Jan-Juil 2013	Mahoumd Sidi Youssef	Amélioration de la segmentation d'images à très haute résolution avec reconnaissance de formes
2ème Ing.	Juin-Août 2012	Mahmoud Sidi Youssef	Classification SVM des propriétés des sols par imagerie hyperspectrale
2ème Ing	Sep-Déc 2004	Mu Xie	Développement d'une plate-forme d'évaluation de MNE (Modèle Numérique d'Élévation)

(***) **L'application de sélection d'attributs (Feature selection application)** [1] **développée au cours de ce stage a été diffusée aux utilisateurs de l'OTB (https://www.orfeo-toolbox.org/) et intégrée le 20 Avril 2015.**

[1] https://github.com/boussaffawalid/FeatureSelection

RAYONNEMENT SCIENTIFIQUE

ECHANGES INTERNATIONAUX

2012 - 2014: Mise en délégation à l'UMR LISAH 144 (IRD, INRA, SupAGro) avec affectation géographique à Tunis. Le projet de recherche a porté sur de nouveaux développements méthodologiques pour la délimitation de parcellaire agricole à partir de données satellitaires à très haute résolution spatiale (GeoEye , Pléiades) et la cartographie de l'occupation du sol à grande échelle en milieu cultivé.

CONFÉRENCES "INVITÉ"

- Semaine thématique en Géospatial, **N. Chehata**,*"Place des drones dans l'offre Très haute résolution spatiale"*, 15 Nov 2016, ENIT, Tunis.
- 1 er Workshop EA G&E Drones pour les sciences de la terre, **N. Chehata et al.**, *"Besoins en imagerie par drone pour les sciences de la terre"*, 11 Mars 2016, ENSEGID Bordeaux-INP.
- Séminaire de recherche à l'EA G&E, **N. Chehata**, *"Drones et capteurs satellitaires à très haute résolution spatiale et spectrale pour les applications en géosciences"*, 28 janvier 2016, ENSEGID Bordeaux-INP.
- Journée ART (Action Régionale Transverse, Aquitaine, pôle Théia, **N. Chehata**. *Présentation des travaux de l'équipe G&E en traitement d'images pour la télédétection*, Bordeaux,12 Nov 2015.
- Colloque drones et moyens légers aéroportés et leurs applications géospatiales en recherche, **N. Chehata**,*"Place des drones dans l'offre Très haute résolution spatiale"*, 03-05 Nov 2015, Tunis.
- Réunion de restitution du projet FORESEE, A. Ferraz, C.Mallet et **N. Chehata**,*"Large-scale road detection in mountaneous areas using airborne lidar data"*, 08-10 Octobre 2014, Nancy.
- Réunion de restitution du projet FORESEE, **N. Chehata** et D. Guyon,*"Potential of multi-source data for forest variable retrieval"*, 08-10 Octobre 2014, Nancy.
- Séminaire GCT 2014 (Geospatial Conference in Tunis),**N. Chehata**, *Potential of remote sensing techniques for mapping and monitoring cultivated landscapes: Application to Lebna catchment, Cap Bon, Tunisia*, 17-21 Mars 2014.
- Séminaire du groupe Lidar et forêts Suisse, C. Mallet et **N. Chehata**, *"Traitement de données Lidar à l'IGN, France"*, 07 Mai 2014, Centre forestier de la montagne de Boudry, Canton de Neuchâtel, Suisse.
- Séminaire de recherche à l'EPFL, C. Mallet et **N. Chehata**, *"Analyse et traitement de donneés Lidar: applications en France"*, 08 Mai 2014, EPFL Lausanne, Canton de Vaud, Suisse.
- Séminaire Pleiades Days, **N. Chehata**, K. Ghariani, M.M. Sidi Youssef, A. Le Bris, C. Mallet, and P. Lagacherie. *"Delineation of anthropic landscape features from VHR satellite imagery: application to agricultural parcel delimitation"*. Toulouse, Apr 2014.
- Séminaire Pleiades Days, B. Beguet, **N. Chehata**, S. Boukir, and D. Guyon. *Quantification and mapping of forest structure from pléiades image texture*, Toulouse, Apr 2014.
- Séminaire les 10 ans Kalidéos, B.Beguet, **N. Chehata**, D.Guyon, and S. Boukir. *"Classification d'images THR de forêt à partir d'attributs texturaux représentatifs des variables forestières"*, Bordeaux, Oct 2012.
- Séminaire GCT 2013 (Geospatial Conference in Tunis), **N. Chehata**, *Apport de la télédétection pour la caractérisation des paysages cultivés: applications sur la région du Cap Bon, Tunisie*, Mars 2013.
- 1 ère journée scientifique des laboratoires de recherche de Sup'Com " TIC et environnement", *"Cartographie numérique des propriétés des sols par imagerie hyperspectrale"*, C. Gomez et **N. Chehata**, 6 Décembre 2012.

- Séminaire Pleiades Days, B. Beguet, **N. Chehata**, S. Boukir, and D. Guyon. *"Detection and mapping of forest changes : application to the maritime pine stands in sw france"*, Pleiades Days, Toulouse, Jan 2012.
- Journée Kalideos, **N. Chehata**, "Caractérisation des changements structuraux des couverts de pins maritimes à l'aide de l'imagerie satellite THR : application à la cartographie des dégâts de la tempête du 24 janvier 2009", 15 Déc 2009, INRA Bordeaux, Villenave d'Ornon.
- Séminaire de recherche EA GHYMAC, **N. Chehata**, *"Lidar et applications environnementales"*,Juin 2007, Bordeaux.
- Séminaire des laboratoires MAP (Modéles et simulations pour l'Architecture, l'urbanisme et le Paysage), **N. Chehata**, *"Etat de l'art des approches automatiques en reconstruction 3D de bâtiments"*, Avril 2002, CNRS UMR 694, Marseille.

ACTIVITÉS D'EXPERTISE

- 2015: Membre **nommé** à la sous-commission "formation et recherche" de la commission Nationale de l'espace Extra atmosphérique, ministère de l'enseignement supérieur et de la recherche de la Tunisie. Créée par le décret n 84-1125 du 24 septembre 1984 (modifié par le décret n 93-1642 du 9 août 1993).
- Depuis 2013: Expertise CNES dans le groupe mission satellitaire hyperspectrale HYPXIM, groupe urbain
- 2009-2013: Expertise CNES dans les groupes d'accompagnement thématiques de pléiades, ORFE0 "GT3 Cartographie" et " GT6 Forêt et agriculture"

ACTIVITÉS ÉDITORIALES

Editeur associé invité 2015

- En collaboration avec Clément Mallet et Grégoire Mercier au numéro spécial" Multitemporal remote sensing change detection ", ISPRS Journal of Photogrammetry and Remote Sensing.

Clément Mallet, Nesrine Chehata, and Grégoire Mercier, Theme Issue "Multitemporal remote sensing data analysis",ISPRS Journal of Photogrammetry and Remote Sensing,107:1–2, 2015.

ACTIVITÉS DE RELECTURE

Depuis 2008 Revues à comité de lecture

- IEEE Transactions on Geoscience and remote Sensing (TGRS)
- International Journal of Photogrammetry and Remote Sensing (IJPRS)
- IEEE Journal of Selected Topics in Applied Earth Observations ans Remote Sensing (JSTARS)
- European Journal of remote sensing (EuJRS)
- Revue Française de Photogrammétrie et de Télédétection (RFPT)
- Journal Forestier Suisse

Depuis 2009 Conférences internationales

- Présidente d'une session poster à ISIVC 2016 (8th International Symposium on Signal, Image, Video and Communications), Tunisie.
- Membre du Comité Scientifique, ISPRS TeanGeo2016, Tunisie.
- Membre du Comité Scientifique, CVPR Workshop MSF 2014 IEEE/ISPRS workshop on Multi-Sensor Fusion for Outdoor Dynamic Scene Understanding

- Membre du Comité Scientifique, co-proposant session Télédétection, 24 ème Rencontre des Sciences de la Terre RST 2014
- Membre du Comité Scientifique, 10 ème et 11ème Atelier Traitement et Analyse de l'Information: Méthodes et Applications, Tunisie (TAIMA'12,TAIMA'13)
- Membre du Comité Scientifique ISPRS – WG III – Photogrammetric Computer Vision (PCV 2010)
- Membre du Comité Scientifique ISPRS – WG III – Photogrammetric Computer Vision And Image Analysis Workshop (LaserScanning'09)

PARTICIPATION AUX RÉSEAUX DE RECHERCHE

Depuis 2016: **participation aux CES Theia**, Centre d'expertise du pôle Théia (pôle thématique surfaces continentales), CES OSO (occupation du sol) et CES Artificialisation- urbanisation.

2009- maintenant, membre **IEEE et GRSS** -Geosciences and remote Sensing society

2016: **participation à la SFPT (Société Française de Photogrammétrie et Télédétection)**, sélection de 3 candidats africains pour la prise en charge de leur participation à la 1 ère école d'été de la SFPT sur « Logiciels Libres pour le Traitement des Images satellites » du 4-8 juillet 2016

2009- maintenant, membre **IEEE et GRSS** -Geosciences and remote Sensing society

2014 - maintenant: **Membre active dans l'ART Aquitaine** (Action Régionale Transverse) du pôle Thématique Surfaces Continentames Théia. Notre rôle est de coordonner les acteurs de télédétection en Aquitaine afin de proposer des projets de recherche à fortes retombées économiques pour la région Aquitaine et notamment dans les secteurs viticoles, forestier et littoral.

2012-2014: **membre active dans l'AST télédétection** (Action Scientifique tranverse) de l'**OASU** (Observatoire Aquitain des Sciences de l'Univers) de Bordeaux, Axe détection et suivi de changements spatio-temporels et multi-échelles de surfaces (continentales et océaniques) par télédétection optique. Montage de projets région en commun avec l'UMR ISPA (INRA), l'UMR IMS et l'UMR EPOC.

2009-2013:**Participation aux groupes thématiques ORFEO du CNES** pour l'accompagnement du lancement du satellite Pléiades, (GT6: Agriculture et Forêt, GT3: Cartographie)

2009-2013:Participation à la **valorisation** des données satellitaires de la **base de données KALIDEOS LITTORAL** du CNES.

2006-2009: Participation active aux réunions du **GDR-ISIS** notamment l'axe CACHANT (Cartographie Automatique des ChAngements Naturels par Télédetection).

JURY DE THÈSE ET DE SÉLECTION

Membre de jury de 3 thèses et 1 comité de sélection pour un poste de maître de conférences
Membre du jury de thèse de Arnaud LE BRIS Décembre 2015

Membre du jury de thèse de Benoit BEGUET Octobre 2014

Membre du jury de sélection de MCU section 61 2012
ENSEEIHT

- Candidat sélectionné: Mathieu Fauvel

Membre du jury de thèse de Li GUO Août 2011

ACTIVITÉS D'ANIMATION SCIENTIFIQUE

- Organisation et coordination d'une journée de veille technologique GEOS2016- Agriculture Numérique, 2 Décembre 2016 à l'INAT. 150 participants. Intervenants académiques et industriels
- Proposition et co-présidence d'une session orale invités "Advanced Methods for Lidar Data Processing", IEEE IGARSS 2017, Fort Worth, Texas,USA.
- Co-organisation du séminaire "Stratégie Nationale Tunisienne en Technologies spatiales", co-animatrice de la session "Infrastructures et Technologie spatiales (Analyse de l'existant, des besoins et des priorités)", en cours.
- Organisation et coordination du 1 er workshop de l'EA G&E Drones pour les sciences de la terre, 11 Mars 2016, Bordeaux, France (40 participants: industriels et académiques)
- Membre du comité d'organisation, colloque drones et moyens légers aéroportés et leurs applications géospatiales en recherche, 03-05 Novembre 2015, Tunis (100 participants).
- Co-présidence de session "Applications, en hydrologie", colloque drones et moyens légers aéroportés et leurs applications géospatiales en recherche.
- Membre du comité d'organisation, GCT 2015 (4 ème édition Geospatial Conference in Tunis), 12-14 Mai 2015, Tunis (300 participants).
- Co-présidence de session "Oral - Image Processing I" , IEEE IGARSS 2013, Melbourne, 2013.

DIFFUSION LOGICIELLE

- Intégration d'une boîte à outils de sélection d'attributs (Feature selection application), dans la librairie OTB (Orfeo Tool Box), développée par le CNES en C++ (sur base ITK) et dédiée au traitement de grands volumes de données de télédétection. Le développement a été réalisé durant le stage de fin d'études de Walid Bouceffa - ENSI Tunisie, en 2014 (cf. liste des encadrements).

PRIX ET DISTINCTIONS

- 2017-2021 : Obtention de la PEDR (Prime d'encadrement doctoral et de Recherche), évaluation au niveau national français par le CNU section 27-Informatique
- La thèse de Benoit Beguet (2011-2014) a reçu le prix THES'AQT (1 ère édition du forum de l'innovation globale NOVAQT), Aquitaine.
- Nomination pour le prix du meilleur article [28]. ICIP 2013 (International Conference on Image Processing),2013.

RESPONSABILITÉS SCIENTIFIQUES

PROJETS DE RECHERCHE PASSÉS ET EN COURS

projet TOSCA CLASS-TIP (40 K€) 2017-2018

- Cartographie du Littoral Aquitain par classification multiple Semi-Supervisée de descripteurs Texturaux extraits d'Images Pléiades
- Coordinateur du projet: EA G&E
- Contribution: **participation au workpackage classification semi-supervisée**

Projet inter-thème EA G&E (2 K€) 2016 - 2017

- Apport des images hyperspectrales pour la cartographie à très haute résolution spatiale pour la cartographie fine de la surface de corps sableux estuariens

- Contribution: **Coordinatrice du projet. Coordination de l'acquisition hyperspectrale par drone en Garonne et dans l'estuaire de la gironde. Mise en place des méthodologies de cartographie des taux de sable/argile et de la cartographie des faciès.**

Projet ANR Mobilité et systèmes urbains - HYEP (500 K€) 2014-2017

- Hyperspectral imagery for environmental urban planning
- Coordinateur du projet: laboratoire LIVE
- **Contribution: participation au WP2: Sélection de bandes utiles pour la caractérisation des matériaux urbains. Encadrement de la thèse d'Arnaud Le Bris. Participation au WP4 en coordination avec les usagers finaux (Toulouse métropole)**.

Projet ANR Transmed ALMIRA (600 K€) 2014-2017

- Adapting Landscape Mosaics of medIterranean Rainfed Agrosystems for a sustainable management of water and soil resources
- Coordinateur du projet: laboratoire UMR LISAH (SupAgro, INRA, IRD)
- **Contribution: Participation au WP1: Cartographie de l'occupation du sol à grande échelle. Coordination des acquisitions d'images satellitaires et des observations des états de surface et l'occupation du sol sur le site du lebna (Tunisie). Coordination de la numérisation du parcellaire sur le Lebna. Encadrement des stages de Walid Bouceffa et Houssem Daagi et de la thèse d'ines ben Slimène**.

Projet PNTS-INSU (12 K€) 2013 - 2014

- Détection d'éléments anthropiques du paysage à impact hydrologique à partir d'images satellitaires à très haute résolution (THR): application à la détection des limites parcellaires dans un milieu cultivé
- Contribution: **Coordinatrice du projet et mise en place de méthodologies de détection de parcellaires**

projet TOSCA STELLA-THR (43 K€) 2011-2014

- Cartographie et suivi des changements dans l'occupation des surfaces des systèmes aquatiques et terrestres par des données THR sur le site atelier aquitain STELLA: préparation à l'utilisation de PLEIADES
  STELLA (réseau Scientifique de TELédétection Aquitain)

- Coordinateur du projet: laboratoire IMS/ ENITAB
- Contribution: **Coordination des applications sur la forêt et co-encadrement de la thèse de Benoit beguet**

Projet Région Aquitaine: STELLA (158 K€) 2011-2014

- Caractérisation de la dynamique spatiale et temporelle des changements majeurs affectant le fonctionnement et la durabilité des écosystèmes sur le site atelier télédétection Aquitain.
  STELLA (réseau Scientifique de TELédétection Aquitain)

- Coordinateur du projet: UR EPHYSE - INRA Bordeaux
- Contribution: **Coordination des applications sur la forêt et co-encadrement de la thèse de Benoit Beguet**

Projet Région Aquitaine iCLEF-THR(49 K€) 2010-2013

- Impact des événements CLimatiques Extrêmes sur les changements Forestiers par télédétection multitemporelle à Très Haute Résolution (THR).

- Contribution: **Coordinatrice du projet** et **co-encadrement de la thèse de Benoit Beguet**
- **labellisation par deux pôles de compétitivités AeroSpace Valley et Xylofutur**

Etude Orféo-CNES GT3 Cartographie (6 K€)) 2010

- Classification orientée objet en milieu urbain pour la mise à jour de bases de données cadastrales à partir de données satellitaires THR.
- **Contribution : Encadrement de stage Master 2**

Etude Orféo-CNES GT6 Forêt (13 K€) 2009-2010

- Détection de changements à partir d'images Formosat pour la cartographie des dégâts de tempête sur la forêt des Landes.
- **Contribution: encadrement de stage de master 2 et d'un CDD**

Projet ANR iTOWNS (181 K€) 2008-2010

iTOWNS: image based Town Online Web Navigation and Search engine
Navigation dans un flux de panoramiques réalisés par des caméras embarquées sur un véhicule, extraction d'informations et indexation pour des requêtes complexes Coordinateur du projet: laboratoire MATIS / IGN **Contribution: Analyse et traitement des nuages de points 3D pour l'extraction d'objets verticaux et encadrement de stage de Master 2**

Projet BQR Bonus Qualité Recherche - Université Bordeaux Montaigne (5 K€) 2009

- Caractérisation des changements structuraux des couverts de pins maritimes : application à la cartographie des dégâts de la tempête Klaus
- **Coordinatrice du projet**
- Financement Université Bordeaux Montaigne
- **Contribution: Préparation en amont de l'étude Orféo-CNES GT6 Forêt et du projet Région Aquitaine iCLEF-THR.**

ADMINISTRATION DE LA SCIENCE

Depuis 2014 : Membre **élue** à la commission recherche de l'ENSEGID

RESPONSABILITÉS COLLECTIVES

Responsabilités à l'ENSEGID Bordeaux INP 2014 - maintenant

- Membre responsable des conventions d'échanges de données avec l'IGN, Géosud et le GIP AtGERI (plateforme PIGMA)
- participation au jury de recrutement sur dossier des ENS2.

Responsabilités à l'UMR LISAH 2013 - 2014
ORE OMERE

- Responsable de la commande des données satellitaires sur le site du Lebna -Tunisie
- Responsable des données d'occupation du sol (terrain et production) de l'ORE OMERE sur le site Lebna

Membre responsable de la gestion des licences logicielles 2008-2011
Institut EGID

Membre co-responsable du site web de l'egid 2007-2011
Institut EGID

RESPONSABILITÉS PÉDAGOGIQUES

Animation pédagogique de la thématique Imagerie- Télédétection-SIG 2015
ENSEGID- Bordeaux INP

· Animation de la thématique sur les 3 années, refonte des enseignements pour favoriser l'intégration avec les applications thématiques pour la prochaine maquette d'enseignement 2017-2021.

Responsabilité des modules d'enseignement en télédétection et géomatique 2014 - maintenant
ENSEGID- Bordeaux INP

· ENS1: SIG et télédétection , ENS2 : Géomatique, ENS3 : SIG et bases de données environnementales

Responsabilité des stages ENS2, ENS3 en télédétection et SIG 2014 - maintenant
ENSEGID- Bordeaux INP

Montage des modules d'enseignement en télédétection et géomatique 2010 -
Nouvelle Maquette enseignement et Obtention du titre d'ingénieur - ENSEGID

· Refocalisation des enseignements d'imagerie en télédétection et développement de la géomatique
· ENS1: SIG et télédétection , ENS2 : Géomatique, ENS3 : SIG et bases de données environnementales

Montage de modules de formation intégrés 2014

Module de 48h : SIG et BD environnementales pour l'élaboration des PPR littoral
réglementation, collecte de données d'enjeux, mise en oeuvre outils de télédétection, SIG, BDG, utilisation de données lidar, interprétation géologique, croisement d'informations

Responsable du parcours Géo-imagerie du master 2 STEE 2009 - 2011
Institut EGID

· master STEE (Sciences de la Terre, Ecologie et Environnement)

Responsable des stages du master 2 STEE, parcours Géo-imagerie 2007 - 2011
Institut EGID

Responsable des modules d'enseignement en traitement d'images et télédétection 2007 - 2011
Institut EGID
Master 1 : Analyse et traitement d'images, Géomatique
Master 2 : Télédétection, Télédétection et environnement

Responsable des projets du parcours Géo-imagerie, Master 2 STEE 2007 - 2011
Institut EGID

RESPONSABILITÉS ADMINISTRATIVES

· **Responsabilités administratives à l'ENSEGID** 2014 - maintenant
ENSEGID- Bordeaux INP

· Membre **nommée** à la commission informatique de l'ENSEGID
· Membre **nommée** responsable des relations internationales avec la Tunisie à l'ENSEGID

Membre élue du CA de la représentation IRD-Tunis 2013 - 2014
IRD Tunis

Membre élue du CA de l'institut EGID 2008-2011
Institut EGID

ACTIVITÉS ASSOCIATIVES

- **Présidente de l'association AGEOS** (African Association for Geospatial Development), depuis Octobre 2015

  - actions de sensibilisation auprès d'étudiants ENIT, Sup'com, FST et INAT
  - organisation d'une formation de photogrammétrie sur logiciel libre MICMAC, Mars 2017, en collaboration avec la SFPT et l'ENSG (Ecole Nationale Des Sciences Géographiques), France.
  - organisation d'une journée de veille technologique GEOS2016- Agriculture Numérique, 2 Décembre 2016, 150 participants, Intervenants académiques et industriels
  - supervision de projets de cartographie : Développement d'un plateforme SIG WEB pour le suivi de l'état de conservation des bâtiments de la médina de Tunis en collaboration avec l'Institut National du Patrimoine.
  - supervision de projets de cartographie : cartographie de 3 circuits de randonnées au 1/25000 dans la région de Siliana, en partenariat avec l'association Tunisienne des Randonneurs (ATR)
- Vice-présidente de l'association AGSI-Tunisia (African Géospatial Sciences Institute) Déc 2013-Juil 2015

Besoins, données et enjeux scientifiques

2

Sommaire

2.1 Besoins en cartographie de l'occupation du sol 33
2.1.1 Caractérisation du milieu urbain 34
2.1.2 Caractérisation du milieu forestier 35
2.1.3 Cartographie des éléments anthropiques en contexte agricole . 36
2.1.4 Cartographie en milieu côtier 38
2.2 Besoins en production de Modèles Numériques de Terrain (MNT) fins 38
2.3 Les données de télédétection utilisées 39
2.3.1 Les images optiques 39
2.3.2 Les données LiDAR 41
2.3.3 Complémentarité des données optiques et LiDAR 43
2.4 Besoins Vs. Données 44
2.5 Les verrous scientifiques 46

Dans ce chapitre, les besoins en caractérisation des différents milieux sont présentés ainsi que les enjeux sociétaux correspondants. Tout au long de mon parcours, différentes thématiques ont été traitées ; l'urbain, le littoral en passant par les milieux forestier et agricole. Les besoins thématiques diffèrent suivant les milieux étudiés, ce qui nous a conduit à utiliser des données multi-sources et multi-capteurs pour y répondre. Ainsi, les données de télédétection optique et LiDAR nécessaires seront présentées en montrant toute la richesse des données optiques en termes de résolution spatiale, spectrale et temporelle et les spécificités des données LiDAR. Nous montrons également la complémentarité des données optique et LiDAR pour certaines thématiques nécessitant la fusion de ces sources de données.

A partir de ces différentes thématiques, j'identifie les verrous scientifiques communs à relever. Le chapitre 3 présentera les principales méthodologies mises en place pour répondre aux verrous identifiés.

Mes travaux de recherches ont répondu à deux types de besoins thématiques ; 1) la cartographie de l'occupation du sol et 2) la production de modèles numériques de terrain fins. Ces produits sont ensuite utilisés pour dériver des indicateurs environnementaux suivant les thématiques correspondantes.

2.1 Besoins en cartographie de l'occupation du sol

Qu'est ce que l'occupation du sol ? L'**occupation du sol** est la description de la couverture physique, naturelle ou anthropologique de la surface de la Terre. La couverture et l'usage des sols sont deux notions différentes et complémentaires pour décrire le territoire.

La **couverture du sol** est une vue physionomique du terrain. Elle correspond à une distinction des éléments sans présager de leur fonction.
Couverture et occupation du sol sont parfois confondues.

L'**usage du sol** est une vue anthropique du territoire suivant sa fonction économique ou humaine. Par exemple, une zone de bâti du point de vue de la couverture du sol pourra être une zone résidentielle, commerciale ou industrielle du point de vue de son usage.

Des bases de données d'occupation du sol ont été constituées à différents niveaux (mondial, européen, national, régional ou local) afin de répondre à différents besoins sociétaux, réglementaires et scientifiques. Une description de l'occupation du sol permet en effet de produire un certain nombre d'indicateurs environnementaux pour la gestion des écosystèmes et des territoires. Une base de données d'occupation du sol, mise à jour, permet de mettre en évidence les évolutions du territoire, d'assurer le suivi d'actions naturelles ou anthropiques et d'en dériver un outil d'aide à la décision pour la mise en place de politiques publiques et l'étude de l'impact des ces politiques.

Par ailleurs, les données d'occupation du sol peuvent aussi être fournies en entrée d'outils de simulation concernant par exemple le climat ou l'hydrologie pour la mise en place de scénarios d'évolutions futures.
En conséquence, les spécifications d'une base de données d'occupation du sol vont dépendre des applications auxquelles elle doit répondre, et de l'échelle à laquelle elle doit être utilisée. Les spécifications d'une occupation du sol vont concerner à la fois :

— sa précision **géométrique**, c'est-à-dire la taille des objets qui vont s'y trouver. On définit ainsi les **unités minimales de collecte** (UMC) qui spécifient pour chaque classe de la nomenclature la taille en deçà de laquelle les objets sont considérés comme trop petits pour apparaître dans la base de données.
— sa précision **sémantique**, c'est-à-dire le niveau de détail de sa nomenclature. Les nomenclatures d'occupation du sol sont le plus souvent hiérarchiques avec un emboîtement allant de classes assez générales à des sous-classes plus fines

Des données d'occupation du sol ont donc été constituées à différentes échelles pour répondre à différents besoins. Certaines sont dédiées à décrire l'intégralité du territoire et ont donc une nomenclature grossière, générique. D'autres sont plus spécialisées, et peuvent se focaliser sur un seul thème

(comme par exemple la forêt, en détaillant les différents types de peuplements dans le cas de la Carte Forestière [1] de l'IGN. Pour le milieu urbain, les différentes bases de données existantes ont été détaillées dans notre ouvrage de livre Briottet et al. [2016].

2.1.1 Caractérisation du milieu urbain

Depuis des décennies, le développement des images satellitaires à haute et très haute résolution spatiale, HR et THR soutient différentes applications en cartographie de l'occupation du sol (ocs) en milieu urbain Briottet et al. [2016]. Parmi les applications les plus importantes, on retrouve :

— la cartographie de l'occupation du sol ;
— la révision cartographique et la mise à jour des bases de données urbaines ;
— la planification urbaine et l'aménagement ;
— l'occupation du sol et sa dynamique d'évolution à travers la modélisation et le développement des scénarios d'évolution.

Plus récemment, d'autres applications considèrent les enjeux du développement durable et de la biodiversité en milieu urbain en étudiant la qualité des écosystèmes, des habitats et des ressources en général. Cette qualité est déterminée par les surfaces et les composants naturels ou semi-naturels (strates végétales, friches urbaines, berges et cours d'eau), mais aussi par les calculs de biomasse, de stockage ou de capture de carbone, de détection de la dégradation des milieux ou de leur fragmentation, etc. L'image est alors complétée par des données ancillaires et/ou utilisée comme variable d'entrée dans divers modèles ou méthodes d'analyse ;

L'imagerie a aussi été récemment sollicitée pour répondre à des enjeux de transition énergétique comme l'évaluation thermique des bâtiments, la localisation de surfaces en toit optimales pour le placement des panneaux solaires ou encore la détection de nouveaux matériaux de construction. La diversité des technologies optiques disponibles, telles que les capteurs THR, hyperspectraux ou infrarouge thermique, fournit une gamme élargie de potentialités pour l'étude du socio-écosystème urbain Briottet et al. [2016].

Cartographie pour la planification et l'aménagement urbains

Des données d'occupation du sol en milieu urbain ont été constituées à différentes échelles pour répondre à différents besoins. Ces données sont produites par photo-interprétation ou bien par des techniques semi-automatiques nécessitant une phase de correction et de validation par photo-interprétation. La description de l'occupation du sol permet de produire un certain nombre d'indicateurs environnementaux pour la gestion des écosystèmes et des territoires, en réponse à différents besoins sociétaux, réglementaires et scientifiques.

En France, l'IGN a entrepris la constitution d'une base de données d'occupation du sol à grande échelle (BD OCS GE) . De précision métrique et compatible avec les autres couches du référentiel à grande échelle (RGE), elle

1. http://inventaire-forestier.ign.fr/spip/IMG/pdf/Guide_technique_cartographie.pdf

a vocation à constituer une nouvelle composante du RGE. L'OCS GE comporte 2 couches, l'une dédiée à la couverture du sol, l'autre à son usage. A chacune de ces composantes est associée une nomenclature hiérarchique. La nomenclature de la couverture du sol comporte 4 niveaux totalisant 14 classes finales. l'unité minimale de collecte varie entre 0,5 ha (en particulier dans les zones construites) et 2,5 ha. La BD OCS GE est produite à la fois par intégration de données existantes et par calcul de nouvelles informations à partir de données existantes, et corrigée et complétée par photo-interprétation.

Cette occupation du sol métrique sera donc bientôt disponible sur la totalité du territoire français. Néanmoins, elle restera insuffisante pour certaines applications en milieu urbain. En effet, l'environnement urbain est un système complexe caractérisé par une forte hétérogénéité de la géométrie des objets qui le constitue (différents volumes et différents niveaux) et des matériaux en présence. Aussi, des connaissances sur la répartition des matériaux en présence sont effectivement nécessaires à certaines applications de modélisation de la ville (simulateurs de micro-météorologie, d'hydrologie, ou de suivi de pollutions), voire de suivi réglementaire (imperméabilisation des sols). La nomenclature de l'occupation du sol doit être plus fine (tant au niveau sémantique qu'au niveau spatial) que la base de données ocs GE. Ainsi, la modélisation fine du milieu urbain nécessite pour certaines applications une connaissance géospatiale des matériaux en présence, autrement dit des cartes de matériaux Le-Bris [2015]. De telles cartes viendraient alors constituer une couche supplémentaire, à la fois dans des bases de données sur l'occupation du sol (comme par exemple l'ocs GE de l'IGN) et dans des maquettes urbaines 3D.

2.1.2 Caractérisation du milieu forestier

L'étude et la surveillance des écosystèmes forestiers, dans un contexte de changements climatiques et de diminution de la biodiversité, est une priorité scientifique, politique et socio-économique très importante à l'échelle régionale en Aquitaine mais aussi nationale.

Les principales menaces pour la forêt sont les incendies, les tempêtes et les défoliations ou mortalités dues aux attaques d'insectes. Les changements climatiques impliquent une pression de plus en plus grande pour les forestiers qui doivent faire face à des catastrophes naturelles maintenant récurrentes comme les tempêtes Martin (1999) et Klaus (2009) qui à elle deux ont ravagé 50% du volume de bois de pin maritime sur pied dans la forêt des Landes de Gascogne, dans le sud ouest de la France. En effet, les tempêtes constituent les perturbations les plus importantes des forêts européennes et sont responsables de plus de la moitié des dégâts de forêts observés durant la seconde moitié du 20ème siècle. Les dégâts de tempête doivent être inventoriés le plus rapidement possible pour prévenir la prolifération des ravageurs et les risques d'incendie. Il est donc nécessaire de localiser les dégâts rapidement et sur de grandes étendues.

L'intensification de ses événements extrêmes affectera fortement le fonctionnement, le développement et la productivité des écosystèmes forestiers. On s'attend en particulier à une modification des aires de répartition de certaines essences forestières et à des dépérissements forestiers ou encore à l'augmentation des risques incendies avec des épisodes de grande sécheresse.

Pour comprendre, évaluer ou prédire ces effets sur des étendues allant du

bassin versant à la région, il est nécessaire de renseigner finement et de quantifier la structure des couverts forestiers et ses changements sous l'effet des pratiques (éclaircies, coupes rases) ou d'événements brutaux tels que les accidents climatiques (sécheresses, tempêtes, ...) ou une agression généralisée par des insectes ravageurs ou encore de changements progressifs dans le temps et l'espace (dépérissements, jaunissement, pertes de feuillage).

L'avènement de l'imagerie à très haute résolution spatiale (THR) a ouvert de nouvelles perspectives pour l'amélioration de la précision géométrique, mais aussi sémantique. Les images à (THR) telles que les images Pléiades (50 cm en panchromatique, 2m en multispectral) rendent possible une description fine de la structure forestière (distribution et dimensions des arbres) à l'échelle du peuplement, en exploitant la relation entre la structure spatiale des arbres et la texture d'image quand la taille du pixel est inférieure à la dimension des arbres. Cette attente répond au besoin d'inventaire spatialisé de la ressource forestière à l'échelle du peuplement et de ses changements dus à la gestion forestière, à l'aménagement du territoire ou aux événements catastrophiques.

Par ailleurs, l'arrivée également des capteurs LiDAR (*cf.* Section 2.3.2) a offert de nouvelles perspectives de par leur capacité à pénétrer le couvert végétal, à restituer avec une grande précision altimétrique le terrain en-dessous, à décrire la structure 3D du couvert et par désormais le déploiement de la technologie LiDAR sur de grandes étendues.

2.1.3 Cartographie des éléments anthropiques en contexte agricole

Les activités anthropiques dans le bassin méditerranéen sont en évolution forte en relation avec une croissance démographique élevée. Dans les zones agricoles, cette croissance entraîne des évolutions considérables de l'occupation du sol. Les pratiques culturales sont modifiées (intensification des productions agricoles dans les terroirs favorables, déprises dans les terroirs non intensifiables). Les paysages sont également remaniés par les aménagements hydrauliques (barrages), hydro-agricoles (retenues collinaires, fossés) et de conservation du milieu (banquettes anti-érosives). Cette activité agricole exerce un impact majeur sur le fonctionnement hydrologique des paysages en modifiant les propriétés hydrodynamiques des sols et en introduisant des discontinuités hydrologiques modifiant les chemins de l'eau. Cette activité produit une série d'éléments anthropiques (états de surface des sols, limites parcellaires, fossés, ...) identifiables à l'échelle locale mais dont l'impact hydrologique se manifeste à une échelle bien plus large, sur plusieurs dizaines à plusieurs centaines de km^2. Dans la perspective de diagnostiquer et de maîtriser ces impacts, il est nécessaire de disposer d'outils de spatialisation de ces éléments anthropiques pour différentes échelles emboîtées de la parcelle au bassin versant.

Certains éléments anthropiques sont partiellement disponibles dans l'espace européen, par exemple les bases de données BDTOPO de l'IGN, le cadastre numérique mais celles-ci restent incomplètes dans l'optique de représenter les impacts hydrologiques. Ceci est particulièrement vrai dans les pays du sud où la disponibilité de données sur les éléments anthropiques est encore plus faible.

Pour répondre à ces enjeux, la télédétection apparaît comme un outil

adapté à l'analyse et la cartographie à grande et large échelle des activités anthropiques des paysages cultivés. Dans ce qui suit, je me focalise sur deux besoins thématiques exprimés en milieu agricole qui sont la délimitation du parcellaire agricole et la cartographie de l'occupation du sol et des états de surface.

Délimitation du parcellaire agricole

En milieu agricole, la parcelle dite 'culturale', unité de production agricole homogène du point de vue du système de culture, s'impose également comme unité hydrologique élémentaire dans la modélisation. Il est donc essentiel de prendre en compte les fortes discontinuités entre parcelles, i.e. de respecter leurs limites. Par ailleurs, les arrangements spatiaux et la connectivité des champs cultivés ont un fort impact sur le ruissellement et l'érosion du sol à l'échelle du bassin versant mais également à l'échelle du paysage.

Dans la perspective de diagnostiquer et de maîtriser ces impacts, il est intéressant de détecter, de façon automatique et à grande échelle, les limites des champs cultivés dans le but de cartographier les types de culture, de suivre les états de surface, de créer ou de mettre à jour une base de données de parcellaire.

Les limites de parcelles peuvent être foncières ou culturales c'est à dire plusieurs cultures dans une même parcelle. Les limites culturales sont dynamiques dans le temps. Une seule image satellitaire mono-date permet d'identifier les limites culturales. A contrario, le parcellaire foncier présente une forte stabilité dans le temps puisqu'il n'est modifié que lors de rares opérations telles qu'une vente ou bien un héritage. L'utilisation d'images satellites multi-dates serait alors plus adaptée pour sa détection.

La détection automatique des limites de parcelles à l'aide d'images satellitaires THR reste un problème difficile en raison d'une forte variabilité des cultures agricoles d'une part mais aussi la variabilité de la nature des limites parcellaires. En pratique, certaines limites peuvent être délimitées par des digues, des levées de terre ou différentes végétations. Une autre difficulté consiste dans la forte variabilité intra-parcelle qui les rend difficiles à segmenter en tant qu'une seule unité homogène.

Cartographie de l'occupation du sol et des états de surface

Les états de surface sont des propriétés des sols déterminantes pour le partage infiltration/ruissellement des surfaces continentales et, en milieu cultivé, ils sont fortement déterminés par les types de productions agricoles et leurs opérations culturales respectives. L'état de surface peut être défini à partir de plusieurs propriétés visibles à l'œil nu à la surface du sol et de leur organisation spatiale bidimensionnelle dénommée Organisation Structurale Superficielle (OSS). Andrieux et al. [2001] propose une typologie des états de surface en fonction des taux d'infiltration pour les cultures méditerranéennes. Les attributs sont : le type de culture, la rugosité du sol, l'encroûtement de la surface, la couverture végétale (taux de recouvrement, herbe, résidus de cultures (litière), et la pierrosité (présence de cailloux).

Dans ce contexte, des données temporelles optiques seront nécessaires afin de suivre l'évolution rapide des états de surface en milieu agricole.

2.1.4 Cartographie en milieu côtier

Les milieux littoraux sont des milieux fragiles sous haute surveillance et il est important de cartographier finement leurs habitats naturels. Leur observation vise à anticiper tant l'évolution de l'occupation du sol parfois en faveur du développement des activités socio-économiques, que les effets du changement climatique, qui peuvent aggraver l'érosion dunaire et augmenter la fréquence et l'intensité des submersions. Ce sont des milieux complexes présentant une biodiversité exceptionnelle qui fait l'objet de nombreuses mesures de conservation et de préservation, d'autant plus prégnantes que ces zones attirent une proportion croissante de la population mondiale.

La côte aquitaine est représentative d'espaces particulièrement sensibles que sont les cordons dunaires, les marais maritimes et prés salés. Afin d'appréhender la complexité de ces milieux et de restituer des indicateurs pertinents liés à la géomorphologie par exemple, mais également pour permettre le suivi d'espèces invasives, la télédétection spatiale, et en particulier les données Pléiades apparaissent comme un support particulièrement adapté. Diverses approches de classification testées sur les données THR ont permis de largement progresser dans l'exhaustivité des faciès pouvant être cartographiés. Les résultats restent toutefois partiels et les approches de classification supervisée trop gourmandes en données de terrain et aussi en images. Les classifications multi-temporelles ou exploitant une grande richesse spectrale restent pour l'instant parmi les stratégies les plus performantes.

L'enjeu est ici de parvenir à une cartographie des indicateurs du littoral à la fois exhaustive et plus précise, notamment en nombre d'espèces ou d'assemblages d'espèces reconnus. Outre le besoin d'améliorer l'approche de cartographie, il faut la rendre plus performante (donc automatisée) et moins coûteuse (en suivi de terrain et images par exemple).

Un autre enjeu sur les milieux côtiers est de détecter rapidement les changements pouvant affecter le littoral au cours, par exemple, d'une tempête hivernale.

2.2 Besoins en production de Modèles Numériques de Terrain (MNT) fins

L'Institut Géographique National (IGN) a pour vocation de décrire la surface du territoire national français et l'occupation de son sol, contribuant ainsi à l'aménagement du territoire, au développement durable, à la protection de l'environnement et à la prévention des risques. Dans un contexte de cartographie de l'occupation du sol pour l'aménagement du territoire, la protection de l'environnement et la prévention des risques, la production de cartographie altimétrique de haute précision devient une donnée majeure pour la réalisation de plans de prévention des risques.

En effet, la connaissance précise des bassins hydrographiques et de leurs zones côtières associées, notamment la hauteur du terrain support des écoulements, mais aussi les éléments obstruant le flux de ruissellement comme la végétation ou certains éléments de sur-sol, est un pré-requis à l'élaboration de cartes des zones inondables et de cartes des risques d'inondation.

Il s'agit ici de la cartographie réglementaire des zones à enjeux, qu'ils soient humains, économiques ou environnementaux. Les données altimétriques is-

sues de la BDAlti® de l'IGN ont une précision hétérogène ($\sigma_z = 2\,\mathrm{m}$ sur 70 % du territoire) et insuffisante en regard de celle des modèles de prédiction du risque hydrologique. Or la cartographie des zones à enjeux nécessite une précision bien plus importante. Elles sont généralement situées dans un environnement topographique varié composé d'un relief parfois irrégulier, de zones de végétation et surtout de zones urbaines.

On appelle **MNT** la représentation mathématique de la surface sol, le plus souvent, sous la forme d'une grille régulière où une valeur d'altitude unique est attribuée à chaque pixel. On s'abstrait donc des objets de sursol qui sont en particulier la végétation et les bâtiments. Le Modèle Numérique de Surface (MNS) regroupe aussi bien le sol que les objets du sursol et décrit ainsi la partie supérieure de la surface, visible depuis une imagerie aéroportée ou satellitaire.

La différence MNS-MNT permet de dériver un produit renseignant directement sur la hauteur des objets du sursol par rapport au sol, appelé Modèle Numérique de Hauteur (MNH). Ce produit est très souvent utilisé pour les applications forestières, pour calculer la hauteur de la canopée. Il permet ensuite d'estimer des variables forestières à l'échelle du peuplement telles que le volume de bois, la biomasse, etc.

2.3 Les données de télédétection utilisées

Afin de répondre aux différents besoins thématiques, des données multi-sources, multi-résolutions ont été utilisées. Dans mes travaux, je me suis focalisée sur les images optiques à très haute résolution (spatiale et spectrale) et les données LiDAR et l'exploitation de leur complémentarité.

2.3.1 Les images optiques

L'intérêt de l'image satellitaire réside dans la stabilité de la qualité des données fournies et la possibilité avec un même flux d'obtenir différentes informations modulées par le nombre de capteurs, le domaine spectral, la résolution spatiale. L'offre actuelle se diversifie en améliorant les caractéristiques techniques pour répondre aux nouveaux enjeux liés à l'anthropisation des milieux. Si l'on considère l'évolution des capteurs satellites optiques Le-Bris [2015], on peut se rendre compte d'une tendance globale vers des données à la fois de mieux en mieux résolues spatialement et de plus en plus riches spectralement.

Jusqu'à récemment, on pouvait distinguer d'une part des satellites d'acquisition d'images à très haute résolution spatiale mais ne comportant que les traditionnelles bandes bleu-vert-rouge-proche infrarouge, et d'autre part des satellites de télédétection (comme par exemple des satellites comme MODIS ou la famille des LandSat) offrant un panel de bandes un peu plus riche mais une résolution spatiale basse ou moyenne.
La fin des années 1990 et le début des années 2000 ont vu l'avènement des satellites à très haute résolution spatiale, avec des capteurs comme GeoEye, QuickBird, Worldview ou Pléiades. Ces satellites permettent de bénéficier de résolution submétrique (de 0,3 à 0,7 m) en panchromatique et métrique (de 1 à 3m) en multispectral. Dans un premier temps, les satellites superspectraux, c'est-à-dire spectralement plus riches, sont restés des satellites à basse ou moyenne résolution spatiale. Il s'agissait par exemple des satellites Landsat (Landsat 7 et 8 comportent respectivement 8 et 11 bandes, avec des résolutions

spatiales de l'ordre de 30 m pour la plupart des bandes) ou MODIS (36 bandes entre 400nm et 14 400 nm, mais avec des résolutions spatiales variant entre 250 et 1000m). Depuis la fin des années 2000 et le début des années 2010, on observe une tendance en faveur d'une nouvelle génération de capteurs spectralement plus riches et de meilleure résolution spatiale. Dans cette mouvance, on peut citer :

— RapidEye (lancés en 2009) qui possèdent une bande dans le red-edge en complément des traditionnelles bandes rouge-vert-bleu-proche infrarouge et acquièrent des images multispectrales de résolution 6,5 m
— Worldview-3 (lancé en 2014) Il s'agit d'un véritable capteur superspectral comportant 8 bandes dans le visible-proche infrarouge et 8 bandes dans le SWIR. Les résolutions spatiales des images sont de 31cm en panchromatique, 1,24m pour les bandes visible - proche infrarouge et 3,7m pour les bandes du SWIR.
— Sentinel-2 (2015). Même si ce satellite est de résolution spatiale moindre (10-20m), il est également un capteur superspectral comportant 12 bandes (6 dans le domaine VNIR et 2 dans le SWIR, mais à une moindre résolution). Avec les satellites Worldview-3 et la mission Venus, on peut réellement parler de capteurs superspectraux à haute voire très haute résolution spatiale.

Par ailleurs, ces nouveaux capteurs satellitaires tendent également à améliorer la couverture spatiale et aussi la résolution temporelle des données avec des temps de revisite de plus en plus courts (tous les 1-5 j).

Les images à haute et très haute résolution spatiale

Pour mes travaux de recherche, je me suis focalisée sur l'étude de l'apport des images optiques HR et THR pour la cartographie de l'occupation du sol dans différents milieux. Nous avons pu utiliser les images aériennes de la BDORTHO de l'IGN mais également des images satellitaires THR type Quickbird ou Pléiades acquises dans le cadre du programme d'accompagnement ORFEO des satellites Pléiades, mené par le CNES. Le CNES nous a également appuyé avec la base de données KALIDEOS LITTORAL sur le bassin d'Arcachon qui nous a permis d'acquérir des images THR et HR sur le site pour diverses thématiques telles que le littoral, la vigne et la forêt. Cette base de données mutualisée a contribué à créer une bonne dynamique régionale en télédétection autour du même site d'études. Dans ce contexte, nous avons utilisé des images Pléiades THR pour la thèse de Benoit Beguet [Beguet, 2014] pour la cartographie et la quantification des structures forestières (projet région Aquitaine iCLEF-THR, et STELLA, projet TOSCA STELLA-THR). Nous avons également pu utiliser des données bi-dates du satellite Formosat-2 pour suivre les dégâts de la tempête Klaus sur le massif des Landes [Chehata et al., 2014b] dans le cadre du projet BQR2009.

A présent, de nouvelles plate-formes telles que GeoSud ou le pôle Théia permettent d'accéder à des données spot et Sentinel gratuites sur la France et sur d'autres régions du monde. C'est dans ce cadre que nous avons acquis les données optiques SPOT6 et 7 pour la thèse d'Ines Ben Slimène sur la cartographie à grande et large échelle de l'occupation du sol en milieu agricole.

Les images à haute résolution spectrale

L'imagerie hyperspectrale consiste à acquérir une image avec un grand nombre (plusieurs centaines) de bandes spectrales étroites et contiguës. De par sa richesse spectrale, cette donnée offre le potentiel d'une caractérisation de l'occupation du sol plus fine qu'en utilisant les données multispectrales traditionnelles à 3 ou 4 bandes. Il n'existe actuellement pas de capteur hyperspectral satellite permettant d'atteindre une résolution spatiale compatible avec l'application de classification des matériaux à très haute résolution spatiale visée. En revanche, il existe actuellement un certain nombre de capteurs hyperspectraux aéroportés, couvrant généralement à la fois le domaine spectral du visible proche infrarouge ou VNIR (allant de 400 nm à 1000 nm) et le domaine infrarouge ondes courtes ou SWIR (allant de 1000 nm à 2500 nm).

Dans nos travaux, nous avons travaillé avec des données aéroportées hyperspectrales dans le cadre de la thèse d'Arnaud Le-Bris [Le-Bris, 2015]. Nous nous sommes intéressés à l'optimisation spectrale d'un capteur superspectral aéroporté, combinant richesse spectrale et bonne résolution spatiale pour répondre à des besoins en occupations du sol spécifiques telles que la caractérisation des matériaux urbains.

2.3.2 Les données LiDAR

Le LiDAR aéroporté

Les systèmes Lidar (Light Detection and Ranging) sont constitués d'un laser, du système synchronisé de réception de la forme d'onde émise et du système de géoréférencement (GPS, centrale inertielle) qui permet de géolocaliser la mesure réalisée. Ces systèmes peuvent être spatiaux, aéroportés (avions, hélicoptères, drones), ou encore terrestres (fixes ou bien embarqués sur un véhicule).

Le laser émet une onde électromagnétique dans une longueur d'onde, suivant une seule direction, on parle de système profileur. Le balayage de la surface terrestre se fait grâce au déplacement du vecteur (satellite, aéroporté, terrestre), conjointement avec un mécanisme de spatialisation, par déviation du faisceau (miroir oscillant ou rotatif, miroir à prisme ou polygone rotatif, fibre optique, ..). Les temps d'aller-retour de l'onde permettent d'avoir une mesure directe des distances séparant le vecteur aérien de la surface terrestre. Le calcul de la trajectographie du vecteur aérien basé sur le couplage d'une unité inertielle et d'un GPS différentiel permet à son tour de transformer ces mesures de distance en points 3D géoréférencées. Les données LiDAR sont donc des nuages de points 3D géoréférencés, non structurés décrivant la surface terrestre.

Ces systèmes présentent deux propriétés majeures : d'abord, la précision altimétrique sur la mesure est particulièrement élevée puisqu'elle est inférieure, sous certaines conditions à 5 cm. En outre, l'émission d'un faisceau lidar au niveau du vecteur aérien permet l'acquisition d'un ou plusieurs points de mesure le long de l'onde réfléchie. Cette propriété est particulièrement intéressante en milieu forestier ou dans les zones de végétation urbaine, où les impulsions LiDAR permettent de traverser la végétation, car elle permet d'accéder non seulement à la hauteur de la canopée, mais aussi, sous certaines

conditions, à l'altitude du sol sous la végétation.

Ces caractéristiques font du LiDAR une technologie adéquate en milieu forestier pour diverses applications telles que la production de MNT fins sous couvert végétal, l'extraction de la desserte forestière ou encore la caractérisation de la structure 3D du couvert végétal (nombre de strates, épaisseur des strates, détection d'arbres individuels,..)

Dans mes travaux de recherche, je me suis intéressée au lidar topographique, permettant de décrire les surfaces continentales. La longueur d'onde utilisée est le proche infrarouge (1064 nm ou 1550 nm correspondant aux fenêtres atmosphériques les plus appropriées). Il existe deux modes d'acquisition des données : 1) le mode "multi-échos" qui enregistre un écho à chaque objet rencontré et le mode "ROC" à retour d'onde complète , ou encore full waveform qui enregistre toute la forme de l'onde retour (*cf.* Figure 2.1). Une forme d'onde représente l'énergie laser rétrodiffusée par la surface terrestre et les objets qui la composent, en fonction du temps et échantillonnée environ toutes les nanosecondes. Les formes d'ondes varient selon les propriétés géométriques et radiométriques des objets illuminés et de leur disposition dans le faisceau laser. Elles dépendent donc de l'empreinte au sol de ce faisceau. Dans mes travaux de recherche, je me suis surtout focalisée sur les LiDAR topographiques aéroportés, à faible empreinte (<1m).

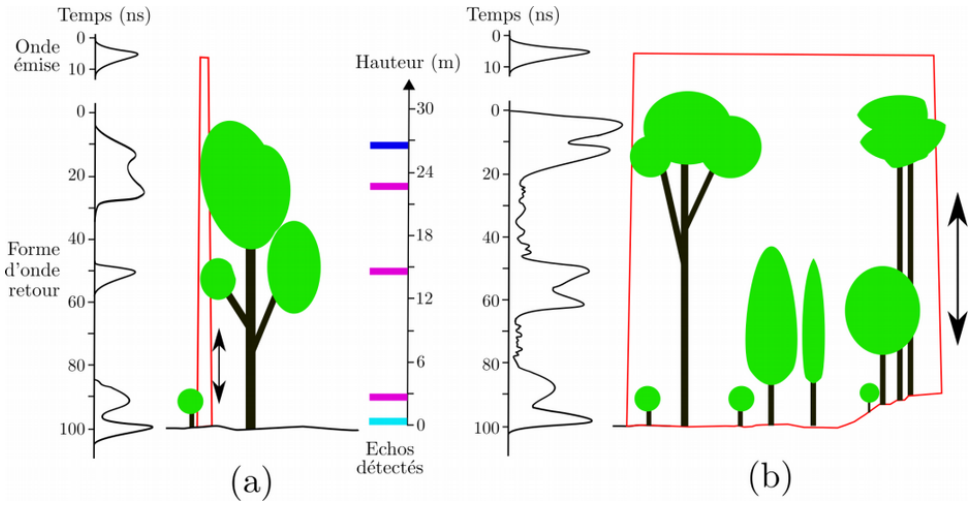


FIGURE 2.1 – *Différentes formes d'onde en fonction de la taille du faisceau laser émis. (a) Capteur aéroporté : chaque objet atteint contribue spécifiquement à la forme d'onde. Un capteur multi-échos se contentera de détecter un faible nombre de distances : le premier écho (bleu foncé), le dernier (bleu clair), voire des échos intermédiaires (rose). (b) Capteur satellite ou large empreinte : la forme d'onde est l'intégration selon l'altitude des objets présents, avec une influence non négligeable du terrain en zones de pente. [Source [Mallet et al., 2016]].*

Nous avons utilisé les données LiDAR multi-échos essentiellement pour la production de MNT fins Bretar and Chehata [2010], la détection de desserte forestière [Ferraz et al., 2016] ou encore la classification en milieu urbain [Chehata et al., 2008] et en milieu côtier [Smeeckaert et al., 2013]. Par ailleurs, nous avons étudié l'apport des données ROC (Full Waveform) pour la classification de scènes urbaines en utilisant de nouveaux attributs en 2D [Chehata et al., 2009b] ou en 3D [Guo et al., 2011].

Le LiDAR terrestre

Comme le LiDAR aéroporté, le LiDAR terrestre est un scanner 3D destiné à imager le sol et le sur-sol à la surface terrestre. Il est composé d'un système avec laser fixe ou embarqué, dans un vecteur terrestre (véhicule), avec un système de spatialisation. Le capteur renvoie des mesures de distance entre le capteur et l'objet rencontré Le véhicule est muni d'un GPS, centrale inertielle et d'un odomètre qui permettent à tout instant de déterminer sa position dans l'espace. Ainsi, en combinant l'information de distance et l'information sur la localisation, il est possible de déterminer la position spatiale réelle de l'objet rencontré.

La spécificité des données terrestres sur véhicule réside dans le fait que la répartition des points 3D dépend directement de la vitesse du véhicule. Une allure rapide permet de couvrir une plus large zone pour un même nombre d'échos (réduction de la densité des points au m^2). Au contraire, une allure lente, voire un arrêt bref du véhicule entraînera une redondance des échos pour une même position spatiale pouvant être gênante pour le traitement du nuage de points.

Nous avons utilisé des données LiDAR terrestre du véhicule STEREOPOLIS dans le cadre du projet ANR iTOWNS, coordonné par l'IGN. Nous nous sommes intéressés à la segmentation de scènes urbaines et tout particulièrement du mobilier urbain tels que les poteaux, et les feux de circulation [Liberge et al., 2010].

2.3.3 Complémentarité des données optiques et LiDAR

En pratique, les différents milieux d'étude sont très complexes, et si l'on s'intéresse à améliorer la précision sémantique des cartographies i.e. utiliser des nomenclatures plus fines, il est nécessaire de combiner les données LiDAR avec d'autres sources de données optiques. En effet, les données LiDAR présentent une forte complémentarité aux données optiques multispectrales ou hyperspectrales.

Les images optiques fournissent une haute résolution spatiale, une information 2D dense sur une grille régulière. Les réflectances spectrales mesurées couvrent un large domaine de longueurs d'onde, du visible à l'infrarouge à courte longueur d'onde (SWIR), en passant par le proche et moyen infrarouge. L'imagerie hyperspectrale fournit une réflectance spectrale beaucoup plus fine que l'imagerie multispectrale, grâce à des centaines de bandes spectrales (une bande spectrale correspondant à une longueur d'onde).

Le LiDAR apporte une information sur la structure 3D et la géométrie des objets, mais fournit un échantillonnage 2D irrégulier et des mesures de réflectances dans le domaine du proche infrarouge ou du vert, respectivement dans le cadre d'un lidar topographique ou bathymétrique. Par ailleurs, le LiDAR a la capacité de pénétrer la végétation jusqu'à atteindre le sol et de renvoyer ainsi des informations sur les régions végétalisées (les routes sous couvert forestier, par exemple, qui sont inaccessibles en utilisant les images optiques seules).

L'extraction d'une surface terrain à partir de techniques photogrammétriques par corrélation automatique peut être mise en défaut d'une part, car le terrain n'est pas visible sous couvert végétal à partir d'une prise de vue

aérienne et, d'autre part, car les zones homogènes en radiométrie et peu texturées ne sont pas adaptées à la recherche de points homologues et répondent mal aux critères d'une bonne corrélation. Dans ce contexte, les données LiDAR apportent une solution adéquate pour le calcul de données altimétriques sous couvert végétal.

La fusion de données hyperspectrales et LiDAR est utilisée en milieu urbain tout d'abord pour la discrimination de différents types de matériaux urbains. Elle l'est ensuite pour exploiter la géométrie 3D des bâtiments et ainsi corriger les réflectances spectrales mesurées par le capteur hyperspectral : cela permet dans le cadre de la classification de l'occupation du sol d'être plus robuste aux conditions d'illumination et aux problèmes d'ombres [Zhang et al., 2013]. Les données hyperspectrales et LiDAR sont également fusionnées en milieu forestier pour l'estimation de variables forestières [Finley et al., 2013] et la discrimination d'espèces végétales [Dalponte et al., 2014]. Une faible densité de points ($\sim$ 4 pts /m^2) permet, à travers la génération d'un modèle numérique de surface, de segmenter les couronnes d'arbres. L'information spectrale est ensuite utilisée pour qualifier les régions obtenues et les classer. Une revue des méthodes de fusion hyperspectrale/lidar pour les écosystèmes forestiers est détaillée dans [Torabzadeh et al., 2014].

Dans mes travaux de recherche, nous avons utilisé conjointement les données multi-sources optique V-R-PIR et LiDAR multi-échos pour la production de MNT fins en milieu montagneux sous couverts forestiers [Bretar and Chehata, 2009] ou encore des données multi-sources optique RVB et LiDAR ROC pour la classification de scènes urbaines [Guo et al., 2011]. Nous avons également mis en évidence la complémentarité des données 3D par photogrammétrie et par LIDAR pour la détection de terrasses en milieu agricole [Sofia et al., 2016].

2.4 Besoins Vs. Données

Les besoins exprimés suivant les milieux d'études peuvent être très différents avec, comme on l'a vu, des spécificités pour chaque type milieu. J'ai essayé de dresser un tableau récapitulatif 2.1 avec différents besoins thématiques et les données nécessaires en face. Je me suis focalisée uniquement sur les données optiques à haute résolution et les données LiDAR. Pour les données optique, j'ai ainsi séparé le mode PAN du mode MS pour les applications qui nécessitent plus une information géométrique ou texturale que spectrale. HSI correspond à l'imagerie hyperspectrale essentiellement utilisée pour des occupations du sol avec une grande précision sémantique. Les données multi-temporelles permettent de détecter et de suivre des changements lents ou abrupts ou encore d'améliorer la discrimination des types de cultures en milieu cultivé. Pour les données 3D, j'ai séparé en trois groupes : données stéréo ou tri-stéréo avec les techniques photogrammétriques ou bien les données LiDAR 3D aéroporté ou terrestre. Les deux dernières sources ont été séparées puisqu'elles répondent à des besoins thématiques différents.

Capteur Milieu	THR PAN	THR MS	HSI	Multitemporel	Stéréo	LiDAR Aéroporté	LiDAR Terrestre
Urbain	extraction d'objets aménagement et planification urbaine	cartographie ocs et usages du sol extraction d'objets	caractérisation des matériaux	détection de changement historique de l'évolution des bâtiments	MNS, MNT extraction d'objets de sursol modélisation 3D de villes	MNS, MNT extraction d'objets de sursol modélisation 3D de villes	modélisation 3D de façades segmentation et indexation de scènes urbaines
Forêt	cartographie structure forestière estimation variables forestières (H, diamètre des couronnes, espacement)	cartographie structure forestière cartographie du type de peuplement estimation variables forestières	différentiation d'espèces végétales cartographie de la santé des forêts	suivi de la santé des forêts suivi des pratiques sylvicoles	délimitation de couronnes volume de bois estimation de variables biophysiques (H, diamètre des couronnes, ...)	production de MNT fins délimitation de couronnes biomasse	modélisation 3D des arbres différentiation d'espèces estimation du diamètre de tige
Agricole	délimitation du parcellaire agricole cartographie des labours	cartographie de l'occupation du sol Niveau 3 (CLC)	cartographie des propriétés pérennes du sol cartographie du stress hydrique et de l'état sanitaire	Cartographie de l'occupation du sol Niveau 4 (CLC) suivi des pratiques agricoles suivi des états de surface extraction de règles d'acteurs	production de MNT fins détection d'infrastructures paysagères suivi de l'érosion ravinaire modélisation hydrologique	production de MNT fins détection d'infrastructures paysagères suivi de l'érosion ravinaire modélisation hydrologique	
Littoral	discrimination de faciès littoraux extraction du trait de côte	discrimination de faciès littoraux extraction du trait de côte	cartographie de dépôts sédimentaires	suivi du trait de côte cartographie de l'occupation du sol Niveau 4 (CLC)	production de MNT fins extraction du trait de côte	production de MNT fins classification eau/non eau	Suivi de falaises suivi glissements de terrain

Table 2.1 – *Exemples de besoins thématiques par type de milieu et les capteurs optique et LiDAR correspondants.*

2.5 Les verrous scientifiques

Les capteurs satellites émergents présentent des hautes résolutions spatiales (entre 1m et 6m) avec de fortes répétitivités temporelles (entre 1 et 5 jours) et une meilleure résolution spectrale. L'outil télédétection est déjà beaucoup sollicité, notamment pour spatialiser et suivre l'évolution dans le temps des occupations du sol. Ces nouvelles données ont créé de nouveaux besoins en cartographie par rapport aux bases de données existantes ; 1) cartographie **à large échelle spatiale** (centaine de km^2), 2) **à grande échelle** (avec des images de haute résolution (1-6m)) et avec 3) une plus **grande précision sémantique** (nomenclature plus fine).

Les enjeux scientifiques à relever, communs aux différentes thématiques, sont à la croisée de différents types d'enjeux qu'on peut relier au 1) contexte thématique appliqué, 2) aux données et 3) aux méthodes d'apprentissage. Ces enjeux sont récapitulés dans le tableau 2.2.

Types d'enjeux	**Verrous**
liés au contexte thématique appliqué	large échelle choix de l'échelle d'analyse gestion des données multi-sources et multi-échelles utilisation de données exogènes opérateur humain dans la boucle
liés aux données	données spatio-temporelles données déséquilibrées données manquantes données bruitées données mal étiquetées peu de données d'apprentissage
liés aux méthodes d'apprentissage	fusion de données multi-sources fusion de données multi-échelles "data shift"

Table 2.2 – *Les différents types d'enjeux et verrous scientifiques à lever en cartographie de l'ocs.*

Quand un besoin thématique est exprimé, le premier verrou qui se pose est l'échelle d'analyse. A quelle échelle le phénomène recherché est-il observable ? Comment choisir cette échelle d'analyse de façon automatique ? Quelles sont les données de télédétection adaptées à cette échelle d'analyse et permettant d'atteindre le niveau de précision sémantique souhaité ?

D'un point de vue appliqué ou "opérationnel", afin de produire une cartographie de l'occupation du sol à **grande** et **large** échelle, nous devons exploiter un grand volume de données satellitaires disponibles, en **multi-capteurs** et en **multi-échelles** mais également utiliser des données exogènes type (MNS, bases de données existantes, ..). Par ailleurs, dans un contexte appliqué, on s'oriente également vers des solutions semi-automatiques avec l'inclusion d'un opérateur humain. En effet, les produits de cartographie obtenus peuvent faire appel à l'opérateur humain en amont pour définir des zones d'apprentissage, spécifier des connaissances a priori, suivant son expertise, dans la chaîne de traitement afin d'introduire des connaissances d'expert ou encore en fin de chaîne pour corriger manuellement certains résultats. Il est donc important de fournir

avec le produit cartographique (ocs ou MNS/MNT) des cartes de confiance qui signalent les zones de forte incertitude nécessitant une intervention humaine. Ce qui permet de réduire considérablement le temps de post-traitements.

Par ailleurs, ces nouveaux capteurs sont de plus en plus riches et posent ainsi d'autres types d'enjeux liés aux données telles que la gestion des données spatio temporelles, la gestion de données manquantes (couverture nuageuse), ou bruitées, ou encore l'existence d'erreurs d'étiquetage si l'on utilise des bases de données qui ne sont plus à jour ou d'une géométrie imparfaite. En pratique, en télédétection, on dispose également de peu d'échantillons d'apprentissage, souvent obtenus par un échantillonnage sur le terrain et ne couvrent pas la variabilité des distributions de classes. Il est donc nécessaire d'adapter le modèle du classifieur pour la cartographie d'ocs à large échelle.

Afin de répondre aux besoins thématiques de cartographie de l'occupation du sol, mes travaux de recherche passés et futurs essaient d'apporter une réponse à certains de ces enjeux qui constituent le fil conducteur de mes recherches. Ainsi, mes travaux se sont focalisés sur les méthodes d'apprentissage automatique. Ces méthodes posent d'autres enjeux qui leur sont liés : Quelles méthodologies pourront êtres mises en place dans un contexte pratique ? classification supervisée, semi-supervisée, apprentissage actif, ...

Pour la cartographie à large échelle, on est souvent confrontés en classification d'images au problème du *datashift*. La distribution des classes est supposée stable sur l'image entière. Les pixels d'apprentissage devraient être représentatifs de l'image entière. Or en pratique, on est confronté à des différences de distributions entre classes à travers l'image et la construction d'un modèle local s'avère sous optimale pour la classification à large échelle. Les méthodologies développées à large échelle devront lever ce type de verrous.

Enfin, pour répondre à un besoin thématique spécifique et afin d'améliorer la précision géométrique et sémantique des produits cartographiques, nous sommes souvent amenés en pratique à utiliser différents capteurs, multi-résolutions, multi-sources. Comment les intégrer dans les méthodes d'apprentissage ? quel modèle de fusion de données utiliser ?

Le chapitre suivant présente un bilan de mes activités de recherche. La méthodologie globale de cartographie de l'occupation du sol en prenant également en compte le cas de données multi-résolutions et multi-sources sera présentée. Dans cette chaîne de traitement, j'ai choisi de me focaliser sur trois axes méthodologiques, pour structurer mes apports méthodologiques.

— Axe 1 : Le calcul et la sélection d'attributs
— Axe 2 : Les méthodes de segmentation 2D et 3D
— Axe 3 : Les méthodes d'apprentissage automatique

Les principales publications seront rattachées aux annexes. Enfin, mes perspectives de recherche seront présentées dans le chapitre 4.

Bilan des activités de recherche

3

Sommaire

3.1 Méthodologies d'élaboration de produits cartographiques par télédétection . . . 51
3.1.1 Méthodologie globale . . . 51
3.1.2 Méthodologie de fusion de données multi-sources . . . 53
3.2 Calcul et sélection d'attributs . . . 55
3.2.1 Calcul d'attributs optiques . . . 56
3.2.2 Calcul d'attributs LiDAR . . . 57
3.2.3 Réduction de dimensionnalité . . . 58
3.2.4 Les critères de sélection d'attributs . . . 60
3.2.5 Les méthodes d'optimisation pour la sélection d'attributs . . . 61
3.2.6 Applications . . . 62
3.3 Segmentation . . . 66
3.3.1 Détection d'objets cartographiques . . . 66
3.3.2 Classification orientée objet . . . 67
3.3.3 Mesures de qualité d'une segmentation . . . 68
3.3.4 Méthodes de segmentation 2D . . . 69
3.3.5 Méthodes de segmentation de données 3D . . . 72
3.3.6 Applications . . . 74
3.4 Classification - Apprentissage automatique . . . 78
3.4.1 Les forêts aléatoires . . . 79
3.4.2 Le concept de marge d'ensemble . . . 79
3.4.3 Optimisation de l'ensemble d'apprentissage . . . 80
3.4.4 Applications . . . 82
Conclusion . . . 89

Ce chapitre présente un bilan de mes activités de recherche passées et actuelles. Mes travaux sont organisés en trois axes méthodologiques : le calcul et la sélection d'attributs, les techniques de segmentation et les techniques d'apprentissage automatique. Dans chaque axe, après la présentation des enjeux scientifiques, les différentes méthodologies proposées seront exposées d'un point de vue théorique et ensuite leurs applications respectives présentées.

Lors de la rédaction de ce chapitre, j'ai fait le choix de me focaliser sur trois axes méthodologiques de la chaîne globale de cartographie de l'ocs : 1) le calcul et la sélection d'attributs, 2) les techniques de segmentation, 3) les techniques d'apprentissage automatique. Dans chaque axe, je présenterai les méthodes proposées, leurs apports, suivis d'une discussion sur les résultats obtenus. Les méthodologies ne seront pas détaillées mais feront référence à une publication, incluse dans le manuscrit en annexes. 5.

3.1 Méthodologies d'élaboration de produits cartographiques par télédétection

Afin de caractériser les milieux naturels, différents besoins émanent des thématiciens afin de valider ou produire leurs modèles, tels que la caractérisation de strates forestières verticales ou la cartographie de la végétation en ville pour des modèles de micro-climat ou encore les limites parcellaires pour la modélisation hydrologique, etc.

La section 3.1.1 présente la méthodologie globale d'élaboration de produits cartographiques, en introduisant les trois axes méthodologiques que je développe dans la suite du chapitre. La section 3.1.2 présente la méthodologie globale de la fusion de données multi-résolutions (pan, MS, HSI, multi-temporelles), multi-sources (optique, LiDAR) et multi-modales (2D et 3D).

3.1.1 Méthodologie globale

Chaque besoin thématique a sa propre échelle d'analyse, échelle à laquelle les objets d'intérêts peuvent être identifiés. Une fois l'échelle d'analyse précisée, les données de télédétection correspondantes sont collectées (*cf.* Table 2.1) ainsi que les données exogènes disponibles. Les bases de données existantes peuvent servir à définir des données d'apprentissage mais également à fournir des données de validation afin de qualifier les produits cartographiques obtenus (*cf.* Figure 3.1).

Pour un utilisateur final (thématicien), non familier à la télédétection, il s'avère très important de lui fournir les attributs les plus pertinents pour son besoin spécifique et de façon automatique. Nous avons ainsi travaillé sur le calcul d'attributs optique et LiDAR novateurs et sur les méthodes de sélection de variables pour choisir de façon automatique les attributs les plus informatifs pour une application donnée. Cette sélection de variables peut être utilisée pour évaluer différents capteurs ou bien pour optimiser les performances d'un classifieur, le paramétrage d'une segmentation ou encore optimiser le jeu de variables explicatives dans une modélisation par régression (*cf.* Section3.2).

Par ailleurs, le thématicien (forestier, agronome, ..) constitue, le plus souvent, sa base de données de référence par des observations sur le terrain et non pas par photo-interprétation. Par conséquent, lui proposer des méthodes automatiques pour optimiser l'échantillonnage terrain et ainsi l'apprentissage du classifieur présente un gain appréciable en termes de temps et de coûts. Ainsi nous travaillons sur des méthodes d'apprentissage automatique visant a améliorer l'ensemble d'apprentissage soit en améliorant le choix des échantillons (*cf.* apprentissage actif Section 3.4.3), soit en rajoutant des informations prove-

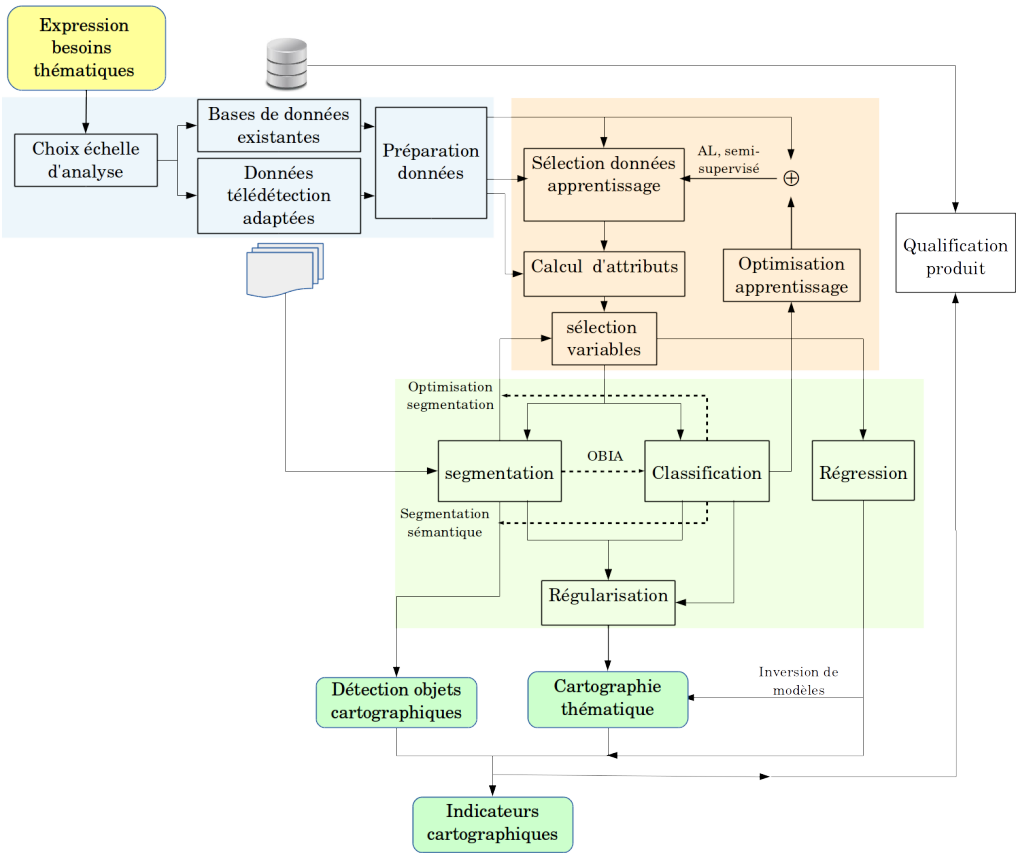


FIGURE 3.1 – *Chaîne méthodologique globale pour l'élaboration de produits cartographiques par télédétection.*

nant de pixels non labellisés, sans avoir un coût supplémentaire de collecte de données sur le terrain (*cf.* apprentissage semi-supervisé Section4.1.1).

Ensuite, trois briques de traitement se détachent : la segmentation, la classification et la régression. La segmentation permettant de partitionner l'images en régions homogènes alors que la classification permet de labelliser des pixels ou des régions. Les méthodes de régressions quant à elles permettent de modéliser des phénomènes continus, i.e. variables biophysiques à partir d'attributs image. Ainsi, la segmentation et la classification vont interagir de plusieurs façons : 1) Tout d'abord par les approches orientées objets (OBIA (Objet Based Image Analysis) (*cf.* Section 3.3.2), 2) les résultats de classifications (labels et probabilités d'appartenance aux classes) peuvent permettre d'optimiser une segmentation (*cf.* Section 3.3.1), 3) la combinaison de résultats de segmentation et de classification permet dans le cadre de la segmentation sémantique d'aboutir à des objets en partant des régions (*cf.* Section 3.3.1). Ceci permet de répondre à un premier besoin thématique sur la détection d'objets cartographiques.

La régularisation est une étape de post-traitement permettant de lisser le résultat de classification en ajoutant des contraintes spatiales et contextuelles avec ou sans segmentation (*cf.* Section 3.3). Cette étape permet d'aboutir au deuxième type de produits cartographiques qui est la cartographie thématique.

Ces deux produits peuvent être ensuite qualifiés et évalués soit en utilisant des données de validation ou bien des bases de données existantes. A partir de ces deux produits, on peut enfin dériver des indicateurs spatiaux, qui

correspondent plus aux besoins de l'utilisateur final (exp, taux de couverture végétale par ilôt, usage du sol, ...).

3.1.2 Méthodologie de fusion de données multi-sources

Dans cette section, je présente un schéma global (*cf.* Figure 3.2) pour la fusion de données multi-résolutions (Pan, MS, HSI, multi-temporelles) mais également multi-sources (optique, LiDAR) et multi-modales (2D, 3D). Le schéma proposé se veut général et toutes les étapes de traitement n'ont pas forcément été abordées dans mes travaux de recherche passés et ne le seront pas dans les perspectives.

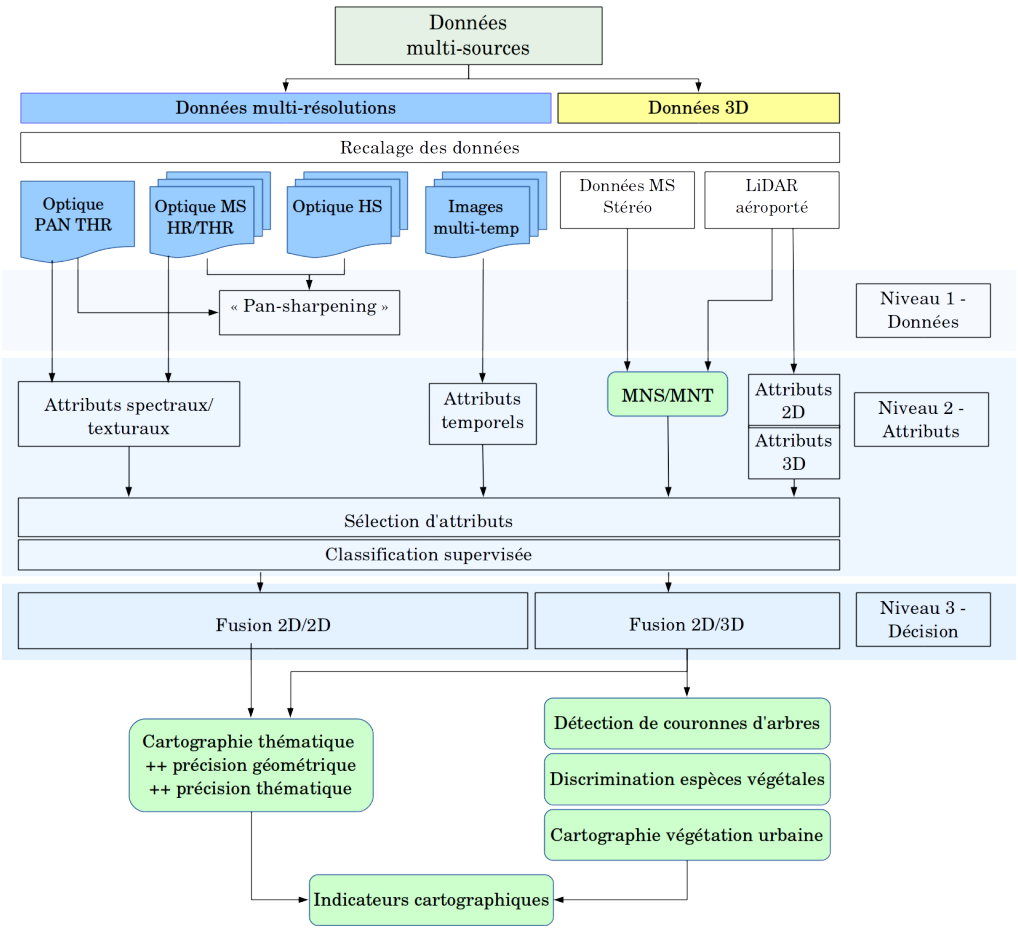


FIGURE 3.2 – *Schéma global de fusion de données multi-sources et multi-résolutions*

La fusion de données multi-sources nécessite un recalage des données. Les données LiDAR sont obtenues par géoréférencement direct, la précision planimétrique est inférieure à 0,5 m. Le géoréférencement des images optiques est plus problématique. Il nécessite d'avoir des points de contrôle au sol et peut induire de grandes erreurs résiduelles en cas de forts reliefs par exemple. Connaissant les paramètres d'orientation externe des capteurs optiques (disponibles avec le système hybride de géoréférencement direct composé d'un GPS et d'une centrale inertielle), deux méthodes existent pour recaler les données LiDAR et optique : 1) "rasterisation" des points LiDAR sur la grille image. Cette méthode consiste à projeter les points LiDAR en 2D et les interpoler afin d'obtenir une image 2D d'altitude ou d'intensité lidar. Cette technique est la

plus souvent utilisée puisqu'elle permet d'appliquer toutes les techniques de traitement d'images 2D. En revanche, elle conduit à une perte d'information et de précision dues aux techniques d'interpolation ; 2) utilisation des points LiDAR 3D en leur affectant les attributs spectraux correspondants, dérivés à partir de l'image optique. La transformation directe terrain- image permet de relier chaque point LiDAR au pixel correspondant dans l'image. Cette technique permet de tirer profit de la structure 3D des données LiDAR, mais uniquement les premiers échos auront des attributs image, car ils seront visibles dans l'image.

La fusion des données peut intervenir à trois niveaux différents :

1. **au niveau des données**, ce qui nécessite une modélisation physique et une calibration du signal (corrections atmosphériques, radiométriques et des effets directionnels). Au niveau des données optiques, la fusion la plus courante est le pan-sharpening entre une image PAN et MS Vivone et al. [2015] ou bien PAN et HSI Loncan et al. [2015]. Le but étant de combiner les détails géométriques de l'image PAN avec une image de meilleure résolution spectrale (MS ou HSI) afin de fournir une image spectrale à une meilleure résolution spatiale. Par contre, cette fusion au niveau données entre optique et LiDAR n'est pas toujours évidente à réaliser et est très peu abordée dans la littérature ; ce premier niveau de fusion n'a pas été abordé dans mes travaux de recherche.
2. **au niveau des attributs** en combinant des attributs multi-sources dérivés de chaque type de données. Pour les données LiDAR, les attributs reliés aux échos, à l'amplitude et à la géométrie des objets pourront être calculés. A partir des données optiques, on pourra notamment dériver des indices spectraux et des attributs texturaux ou encore des attributs temporels. Des méthodes de sélection de variables pour la classification permettent ensuite de garder les meilleurs attributs pour chaque classe. Dans nos travaux [Guo et al., 2011, Chehata et al., 2009b], nous avons montré l'apport de la fusion de données lidar à retour d'onde complète et multispectrales pour la classification urbaine. Dans le premier cas, la fusion est faite en affectant chaque point LiDAR 3D par les attributs multispectraux correspondants, dans le deuxième cas, les points LiDAR ont été projetés et interpolés en 2D.
3. **au niveau des décisions**, où chaque source de données est traitée séparément et ensuite les produits sont fusionnés au niveau décisionnel en utilisant des techniques de fusion bayésienne [Duda and Hart, 1973] ou par la théorie de Dempster-Shafer [Poulain et al., 2011]. Un panorama méthodologique sur les méthodes de fusion de données est présenté dans [Zhang, 2010].
   Dans la littérature, les données LiDAR sont le plus souvent fusionnées à des données hyperspectrales. Les données hyperspectrales fournissent des réflectances spectrales très fines, mesurées par longueur d'onde et permettent donc de discriminer finement des classes ayant des structures 3D similaires mais des caractéristiques spectrales différentes Dalponte et al. [2008], Khodadadzadeh et al. [2015].

Au niveau des produits cartographiques, les données stéréo (ou tri-stéréo) (par techniques photogrammétriques) ainsi que les données LiDAR 3D per-

mettent d'obtenir des MNS et/ou des MNT. Le passage du MNS vers le MNT nécessite des étapes de filtrage des points LiDAR 3D qui n'ont pas été exposés dans ce schéma. Ça sera montré comme application dans la section 3.4.4.

La fusion de données permet d'améliorer la précision géométrique et sémantique des produits cartographiques. Les fusions 2D/2D portent sur des données optiques multi-résolutions ou bien des données multi-sources optique/LiDAR où les données LiDAR auront été interpolées en 2D. Alors que les fusions 2D/3D exploitent les attributs LiDAR dérivés en 3D et la géométrie du nuage de points. Elles sont souvent plus complexes à mettre en place et sont principalement utilisées dans le cas où l'on a besoin d'exploiter la capacité de pénétration du LiDAR à travers la végétation pour caractériser la structure de la végétation. Le principal apport des données LiDAR en fusion de données portera par conséquent sur les produits en lien avec la végétation pour la détection de couronnes d'arbres ou bien la cartographie de la végétation urbaine, ou encore la discrimination d'espèces végétales.

3.2 Calcul et sélection d'attributs

L'augmentation du nombre de capteurs satellitaires multi-sources et multi-échelles avec de meilleurs résolutions spectrales génère de grands volumes de données contenant souvent de l'information redondante.

Cette forte dimensionnalité des données peut générer un certain nombre de difficultés. En premier lieu, le phénomène bien connu de Hughes [Hughes, 1968] risque de se produire lorsque l'on va chercher à procéder à une classification de ces données. Toutefois, des classifieurs robustes tels que les Séparateurs à Vaste Marge (SVM) ou les Forêts aléatoires (ou Random Forests (RF)) sont moins sensibles à ces problèmes [Melgani and Bruzzone, 2004, Camps-Valls and Bruzzone, 2005] et parviennent malgré tout à produire de bons résultats, sauf dans des cas où très peu de données d'apprentissage sont disponibles [Pal and Foody, 2010].
En second lieu, les temps de calcul requis pour traiter ces grands volumes de données vont également être importants. En troisième lieu, se posera la question du stockage de ces données.

Il va donc souvent s'avérer intéressant de réduire la dimension de ces données. Cela va consister à réduire le volume de données (i.e. le nombre de bandes ou d'attributs) de manière à optimiser un certain critère, comme par exemple minimiser la perte d'information utile, et en particulier dans le contexte d'une classification supervisée, de s'efforcer de conserver la meilleure séparabilité, la meilleure distinction entre les classes.

Le calcul et la sélection d'attributs constitue ainsi l'étape préliminaire de toute analyse et traitement de données de télédétection (segmentation, classification) où l'on commence par calculer des descripteurs ou attributs décrivant les pixels ou régions de l'image ou encore les points 3D LiDAR. Dans un deuxième temps, des méthodes de sélection automatique d'attributs permettent d'optimiser le choix de ces attributs en fonction d'un critère donné, par exemple maximiser la précision globale de classification de l'occupation du sol.

Par ailleurs, la sélection d'attributs est indispensable dans le cas de fusion de données multi-sources au niveau des attributs. Ainsi des attributs spéci-

fiques à chaque source de données sont calculés et les méthodes de sélection permettent de ne garder que les attributs les plus pertinents et de prioriser ainsi les sources de données. Les méthodes de sélection d'attributs permettent aussi d'évaluer et comparer des capteurs et leur importance pour une thématique donnée. Si la sélection est appliquée à des bandes spectrales, elle permet également d'optimiser la configuration spectrale de capteurs imageurs pour une thématique spécifique [Kandasamy et al., 2009, Le-Bris, 2015].

Dans cet axe de recherche, j'ai essentiellement travaillé sur le calcul d'attributs optiques et LiDAR où nous avons proposé des attributs LiDAR innovants en multi-échos. Nous nous sommes ensuite intéressés aux techniques de sélection d'attributs pour répondre à des problématiques spécifiques. Deux thèses que j'ai co-encadrées comme encadrante principale ont porté sur l'axe sélection d'attributs. Tout d'abord, la thèse de Benoit Beguet [Beguet, 2014] en collaboration avec l'UMR ISPA (INRA-Bordeaux), a porté sur la caractérisation de la structure forestière par imagerie satellitaire THR. L'objectif était de montrer l'intérêt de données de texture THR multi-échelles pour la caractérisation du milieu forestier. La seconde thèse est celle d'Arnaud Le Bris [Le-Bris, 2015] en collaboration avec l'équipe MATIS du laboratoire LasTIG de l'IGN sur la sélection de bandes utiles à partir de données hyperspectrales pour la caractérisation des matériaux urbains. La finalité de ce travail était d'optimiser la configuration spectrale (position et largeur des bandes) d'un capteur superspectral.

Par ailleurs d'autres travaux ont porté sur l'utilisation des mesures d'importance de variables des forêts aléatoires pour montrer l'intérêt thématique de certaines bandes ou indices spectraux et texturaux, par exemple pour la cartographie des labours en milieu cultivé [Chehata et al., 2013], ou la délimitation du parcellaire agricole [Chehata et al., 2015]. Nous avons également utilisé la sélection d'attributs lors de la fusion de données optiques et LiDAR ROC au niveau des attributs. Nous avons ainsi montré l'importance de certains attributs LiDAR ROC pour la classification urbaine et tout particulièrement l'apport de données multi-sources optique et LiDAR pour la cartographie en milieu urbain [Guo et al., 2011].

Concernant la comparaison de capteurs, suite à une demande des services de production de l'IGN, nous avons fait l'étude de l'apport de données satellitaires RapidEye (disposant d'une bande spectrale dans le red-edge) en comparaison à des données aériennes RVB de la BDOrtho pour la classification de zones forestières en trois classes : feuillus, résineux et mixtes [Le-Bris et al., 2013]. Plusieurs attributs spectraux et texturaux ont été dérivés. Différentes méthodes de sélection d'attributs ont été testées avec trois classifieurs : Maximum de vraisemblance, SVM et les random forests. Cette étude a permis de conclure que le capteur RapidEye apportait surtout l'information spectrale alors que les images THR apportaient un plus grâce aux attributs texturaux.

3.2.1 Calcul d'attributs optiques

En imagerie optique, on utilise rarement les bandes initiales de l'image multispectrale, mais plutôt des canaux transformés calculés à partir des bandes initiales (rouge - vert - bleu - proche infrarouge) de l'image. Ces canaux dérivés permettent souvent de mieux discriminer certains paramètres, à l'exemple de

l'indice spectral NDVI (Normalized Difference Vegetation Index) utilisé pour détecter la végétation. Ces canaux dérivés, ou néo-canaux, appelés aussi attributs en cas de classification, peuvent être de plusieurs types :

— Attributs spectraux : ce sont des indices spectraux tels que les indices de végétation, les indices de brillance, les indices d'eau, ou bien des canaux transformés dans un autre espace colorimétrique (espaces intensité/teinte/saturation ou Lab par exemple), ou encore tous autres indices obtenus par simple combinaison des canaux initiaux ;

— Attributs de texture : ils sont calculés au niveau du pixel (en intégrant l'information au sein d'un voisinage donné) ou de la région (si l'on dispose d'une segmentation) et caractérisent l'information texturale (structure, rugosité, contraste, etc.) de l'image. On pourra par exemple citer les indices texturaux de Haralick Haralick et al. [1973] ou de Gabor [Grigorescu et al., 2002], les profils morphologiques d'attributs [Mura et al., 2010] ou des approches basées sur des descripteurs comme les histogrammes d'orientation du gradient (Histogram of Oriented Gradient ou HoG) [Dalal and Triggs, 2005] ;

— Attributs objet : il s'agit d'attributs spécifiques aux approches objets, basés sur leurs géométries.
L'utilisation de canaux dérivés adaptés ou de leur combinaison permet souvent d'améliorer les résultats de la classification et de mieux discriminer certaines classes. Le bon choix de ces attributs est primordial et conditionne les précisions de classification de l'occupation du sol.

3.2.2 Calcul d'attributs LiDAR

Afin d'analyser un nuage de points 3D, il faut qualifier chaque point LiDAR 3D par un ensemble d'attributs qui va permettre ensuite de segmenter le nuage de points en régions homogènes, de reconnaître des formes prédéfinies ou encore de classer chaque point 3D ou chaque région 3D, et de lui affecter un label prédéfini dans une nomenclature donnée.

Chaque point 3D est qualifié par un certain nombre d'attributs, calculés la plupart du temps en 3D, en analysant la distribution des points 3D voisins ayant les mêmes caractéristiques. Les voisins sont obtenus soit par une requête de type "Plus Proches Voisins" soit par inclusion dans un voisinage de forme prédéfinie centré sur le point d'intérêt : les plus courantes sont cylindrique (sans contrainte sur l'altitude des points voisins, elle devient une requête en 2D, utile pour le calcul d'attributs fondés sur des histogrammes ou des différences d'altitude) ou encore sphérique si l'on ne cherche qu'à récupérer tous les points dans un rayon fixé [Mallet et al., 2016]. La forme peut être également choisie après analyse de la distribution des points 3D dans un voisinage cylindrique par défaut. Une analyse en composante principale permet de déterminer si le voisinage est plutôt sphérique, surfacique ou linéaire. La taille du voisinage est le seul paramètre à fixer et elle est contingentée par la densité de points 3D et l'échelle d'analyse souhaitée. Les voisinages fixes sont plus faciles à mettre en place mais ne sont pas adaptés à différents contextes et souffrent du mélange d'objets : les attributs géométriques qui en découlent ne seront pas optimaux. Les voisinages adaptatifs [Filin and Pfeifer, 2005, Demantke et al., 2011] prennent, par contre, en compte la distribution des points 3D assurant

ainsi la fiabilité des voisinages 3D (au sens d'un critère géométrique prédéfini comme la planarité ou l'orientation des normales) et un meilleur calcul des attributs 3D [Unnikrishnan et al., 2010]. Cependant, ils nécessitent un temps de calcul plus élevé.

En pratique et pour le cas particulier de la fusion avec des données optiques, les attributs de points 3D sont souvent interpolés sur une grille régulière en 2D afin de pouvoir gérer de grands volumes de données et la forte dimensionnalité (nombre d'attributs) des points 3D. Les attributs LiDAR multi-échos (i.e., issus des points 3D) peuvent être classés en 4 familles : ils peuvent être fondés sur (1) les altitudes ; (2) les rétrodiffusions multiples (échos) ; (3) la géométrie 3D locale ou (4) l'intensité lidar. On peut y ajouter la famille d'attributs spécifiques dérivés des formes d'onde complètes 3.3.

Tous ces attributs ont été proposés et détaillés dans notre travail [Chehata et al., 2009b], annexé au document. Par ailleurs d'autres attributs LiDAR 3D, basés sur la densité de points 3D, ont été proposés pour la classification des zones côtières en zone d'eau vs. non eau [Smeeckaert et al., 2013].

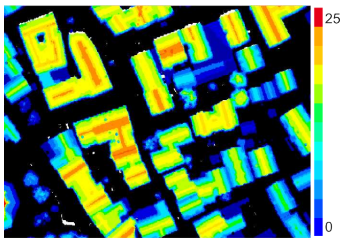
3.2.3 Réduction de dimensionnalité

Le calcul de différents attributs spectraux, texturaux et géométriques, la fusion de données multi-sources ou encore les données de télédétection hyperspectrales génèrent des volumes de données importants. Les méthodes de réduction de la dimension peuvent être distinguées en 2 catégories : les méthodes d'**extraction d'attributs** (*feature extraction*) et celles de **sélection d'attributs**(feature selection).

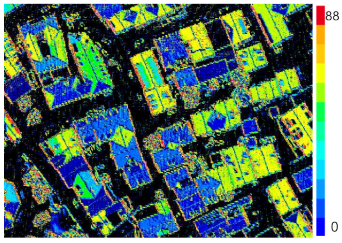
Les méthodes d'extraction d'attributs consistent à résumer l'information contenue dans le jeu de données initial, par exemple en reprojetant les bandes initiales sur des axes choisis de manière optimale. Parmi les algorithmes d'extraction d'attributs les plus connus, on peut citer l'Analyse en Composantes Principales (ACP), le *Minimum Noise Fraction* (MNF) ou l'Analyse en Composantes Indépendantes (ACI) pour le cas non supervisé ainsi que l'Analyse Discriminante Linéaire, ou Analyse Discriminante de Fisher pour le cas supervisé.
A l'inverse, les méthodes de sélection d'attributs vont choisir un sous-ensemble d'attributs issus du jeu de données initial, c'est-à-dire ne conserver que les attributs les plus utiles i.e. les plus pertinents, au sens d'un certain critère. Cette seconde famille permet de conserver le sens physique des bandes qui sont sélectionnées.

Nos travaux de recherche se sont focalisés sur les méthodes de sélection d'attributs. La sélection d'attributs fait partie de la famille des problèmes d'optimisation. En effet, ce travail va consister à sélectionner le sous-ensemble d'attributs permettant d'optimiser un certain critère. Un algorithme de sélection d'attributs est donc l'**association** d'**un critère de sélection** que l'on va chercher à optimiser et d'**une méthode d'optimisation**. Ce critère de sélection d'attributs est une mesure/un score mesurant la pertinence du sous-ensemble d'attributs sélectionnés [Le-Bris, 2015].

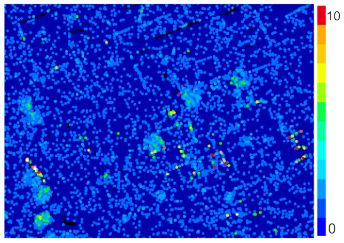
(a) Orthoimage RVB.



(b) LiDAR ME- Δz.

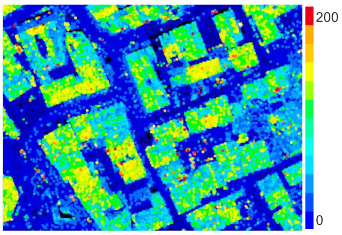


(c) LiDAR FW - angle de déviation N_z.

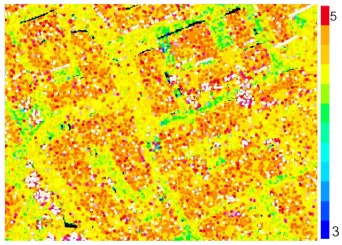


(d) LiDAR ME - Nb échos N.

(e) LiDAR ME - Nb échos normalisé N_e.



(f) LiDAR FW - Amplitude A.



(g) LiDAR FW - Largeur écho w.

FIGURE 3.3 – *Exemples d'attributs LiDAR multi-échos et à retour d'onde complète.*

3.2.4 Les critères de sélection d'attributs

Les méthodes de sélection d'attributs (et en particulier le critère de sélection) sont souvent distinguées en 3 catégories : filtres, *wrapper* and imbriquées (*embedded*). La thèse d'Arnaud Le Bris [Le-Bris, 2015] présente une revue détaillée de ces méthodes ainsi qu'un récapitulatif d'un certain nombre d'approches de la littérature triées selon ces catégories.

Les méthodes *filtres*

Les méthodes de type filtre utilisent un critère de sélection d'attributs indépendant de tout classifieur.
Certaines méthodes vont calculer un score d'importance individuelle par attribut et n'en conservent que les n meilleurs. Ces scores peuvent être supervisés, utilisant la vérité terrain, tel que l'algorithme ReliefF [Kononenko et al., 1997], le score de Fisher [Hall and Holmes, 2003] ou non supervisés (utilisant une décomposition par ACP par exemple) [Chang et al., 1999, Kandasamy et al., 2009].

Cependant, ces mesures d'importance ne prenant pas en compte les dépendances entre attributs, les n meilleurs attributs ne constituent pas un sous-ensemble optimal d'un point de vue de la parcimonie [Le-Bris, 2015]. Dans cette famille *filtre*, d'autres méthodes calculent des scores de pertinence pour des sous-ensembles d'attributs en non supervisé [Chang and Wang, 2006] en gardant un sous-ensemble d'attributs les moins corrélés entre eux. Dans le cas supervisé, des mesures de séparabilité comme celle de Fisher [Fisher, 1936], Bhattacharya ou Jeffries-Matusita (JM) permettent de mesurer la capacité d'un ensemble d'attributs à discriminer différentes classes [Bruzzone and Serpico, 2000, Herold et al., 2003, Serpico and Moser, 2007].

Enfin des mesures statistiques issues de la théorie de l'information, comme la divergence, l'entropie ou l'information mutuelle, sont aussi utilisées pour sélectionner les meilleurs sous-ensembles d'attributs, à la fois en non supervisé [Martínez-Usó et al., 2007, Le Moan et al., 2011] et en supervisé comme dans [Battiti, 1994, Guo et al., 2008, Sotoca and Filiberto, 2010, Cang and Hongnian, 2012].

Les méthodes *Wrappers*

Cette famille de sélection mesure la pertinence d'un sous ensemble d'attributs avec la qualité de la classification correspondante. Une méthode *wrapper* sera donc associée à un classifieur, comme par exemple les Séparateurs à Vaste Marge (SVM) [Li et al., 2011, Zhuo et al., 2008], un classifieur par maximum de vraisemblance [Zhang et al., 2007, Fauvel et al., 2014], les forêts aléatoires (random forests) [Díaz-Uriarte and De Andres, 2006, Guo et al., 2011].

Les méthodes de sélection type *wrapper* garantissent généralement une bonne parcimonie pour de bonnes performances de classification. Néanmoins, ces solutions peuvent être trop dépendantes du classifieur et les temps de calcul sont aussi généralement plus importants [Le-Bris, 2015].

Les méthodes *Embedded*

Les méthodes de sélection d'attributs de type "imbriqué" ou "*embedded*" sont également liées à un classifieur, mais le score mesurant la pertinence d'un sous-ensemble d'attributs sera différent d'une simple mesure de la qualité de la classification. En général, avec ce type d'approches, la sélection d'attributs est effectuée directement lors de la phase d'apprentissage.

Certaines approches se basent sur des modèles de régularisation : un classifieur est entraîné de manière à optimiser une fonction objectif qui comprend un terme d'attache aux données (minimisant le taux d'erreur de classification) et un terme de régularisation qui va privilégier les modèles L1-SVM [Zhu et al., 2004] ou LASSO (Least Absolute Shrinkage and Selection Operator) [Tishirani, 1996]. Ces approches sont efficaces et rapides. En revanche, critère de sélection d'attributs et méthode d'optimisation y sont également complètement liés.

D'autres classifieurs sélectionnent de part la nature de leur méthode d'apprentissage les attributs les plus pertinents. Ainsi, pour les Random Forests (RF) [Breiman, 2001] par exemple, seuls les attributs les plus discriminants (par exemple au sens du critère d'impureté de Gini) sont utilisés lors du découpage d'un nœud. Cette sélection d'attributs à la volée élimine les attributs les moins utiles mais n'est pas pour autant parcimonieuse. Outre cette sélection d'attributs utiles lors de l'apprentissage, les Random Forests fournissent également une mesure d'importance des différents attributs au sein de son modèle [Breiman, 2001]. Cette importance est calculée pour chaque attribut à partir des échantillons *Out-Of-Bag* ou OOB, c'est-à-dire non pris en compte pour la construction d'un arbre de la forêt. Ainsi, pour mesurer l'importance d'un attribut donné, ses valeurs sont permutés aléatoirement pour tous les échantillons OOB. On calcule alors la différence entre le taux de bonne classification de ces échantillons OOB avant et après cette permutation. L'importance de l'attribut est alors estimée par la moyenne des différences sur tous les arbres de la forêt. L'importance individuelle RF d'un attribut prend en compte les autres bandes de manière sous-jacente. Toutefois, l'importance d'une bande aura tendance à diminuer si celle-ci est très corrélée avec de nombreuses autres bandes [Genuer et al., 2010].

L'algorithme SVM-RFE (SVM - Recursive Feature Elimination) proposé par [Guyon et al., 2002] élague récursivement les attributs les moins importants pour un modèle SVM. Cette approche SVM-RFE lie complètement le score de pertinence des attributs et le processus d'optimisation.

3.2.5 Les méthodes d'optimisation pour la sélection d'attributs

Une fois le critère à optimiser choisi, le problème consiste à optimiser le sous-ensemble d'attributs au sens de ce critère. Ces méthodes d'optimisation peuvent être généralement distinguées en méthodes séquentielles ou incrémentales et méthodes stochastiques.

Stratégies séquentielles ou incrémentales

Ces méthodes vont consister à progressivement rajouter ou retirer des attributs au sous-ensemble d'attributs sélectionnés [Pudil et al., 1994]. Par exemple,

la méthode Sequential Forward Search (SFS) commence par sélectionner un attribut puis ajoute progressivement à la sélection un nouvel attribut permettant d'optimiser le critère de sélection. A l'inverse, la méthode Sequential Backward Search (SBS) est initialisée par l'ensemble des attributs puis supprime progressivement de la sélection le plus mauvais attribut au sens du critère de sélection. Plusieurs variantes existent telles que les algorithmes "flottants" Sequential Forward Floating Search (SFFS) ou Sequential Backward Floating Search (SBFS)[Pudil et al., 1994] qui autorisent la remise en question de la sélection d'attributs obtenue au niveau précédent.

Algorithmes stochastiques

Les algorithmes stochastiques font intervenir le hasard dans leur exploration de l'espace des solutions. Ce caractère aléatoire au niveau de l'initialisation et de la recherche d'une solution permet de proposer différentes solutions de qualité équivalente à partir d'un même jeu de données [Le-Bris, 2015]. Plusieurs algorithmes d'optimisation stochastique ont été mis en œuvre pour la sélection d'attributs tels que les les algorithmes génétiques (AG)[Estévez et al., 2009, Zhuo et al., 2008, Minet et al., 2010, Li et al., 2011], l'algorithme Particle Swarm Optimization (PSO) [Yang et al., 2007, 2012], l'algorithme des colonies de fourmis [Zhou et al., 2009] ou encore le recuit simulé [De Backer et al., 2005, Chang et al., 2011].

Parmi ces différentes heuristiques, les algorithmes génétiques et PSO sont parmi les plus utilisés. Les algorithmes génétiques présentent l'avantage de pouvoir prendre en compte facilement des contraintes complémentaires pour guider la recherche d'une solution [Estévez et al., 2009].

3.2.6 Applications

Les méthodes de sélection d'attributs ont été appliquées dans mes travaux de recherche sur différentes thématiques en choisissant à chaque fois les méthodes les plus appropriées. Trois applications phares de mes travaux seront détaillées et les publications correspondantes annexées. D'autres utilisations seront également évoquées.

Modélisation de la structure forestière

Dans la thèse Benoit Beguet [Beguet, 2014] et en collaboration avec l'UMR ISPA (INRA Bordeaux), nous nous sommes intéressés à la quantification des structures forestières à travers l'estimation de variables biophysiques à partir d'images THR, de type Pléiades.

L'objectif de ce travail a été d'évaluer une méthode générique d'estimation des variables forestières utiles à inventaire de ressources forestières en exploitant au mieux le potentiel de la texture d'images THR, multi-résolutions (panchromatique et multispectral). Ces variables forestières sont : la hauteur des arbres, le diamètre des couronnes, la densité, le diamètre des troncs et l'espacement entre les arbres. Pour cela, nous avons utilisé les matrices de co-occurrence pour analyser la texture [Haralick et al., 1973]. Afin de pallier au problème du paramétrage spatial des attributs de texture issus des matrices de co-occurrence, l'idée a été de ne pas faire d'hypothèses a priori sur le paramétrage spatial et de calculer un très grand nombre d'attributs de texture en

balayant une vaste gamme de descripteurs multi-résolutions (pan et MS) mais aussi multi-échelles (paramétrages différents de l'orientation, le déplacement et la fenêtre d'analyse).

L'estimation des variables forestières a été réalisée par modélisation statistique par régressions linéaires simple et multiple [Beguet et al., 2012, 2014].

Pour la régression linéaire multiple, nous cherchions à développer des modèles parcimonieux (avec un nombre limité d'attributs image comme variables prédictives) tout en réduisant au minimum la colinéarité au sein de ces modèles, cela de façon automatique. Les développements méthodologiques proposés sont donc l'automatisation de la sélection du modèle le plus performant, ainsi que l'automatisation de la sélection de combinaisons optimales des résolutions spatiales et bandes spectrales disponibles. Pour la méthode de sélection de variables en cas de régression multiple, nous avons proposé une nouvelle méthode stochastique, nommée Random PRiF adaptée au problème de la colinéarité. Les performances de cette nouvelle méthode ont été comparées à des méthodes classiques séquentielles ou encore des approches embedded par random Forests. Les travaux sont détaillés dans la publication [Beguet et al., 2014] en annexes, section 5.1.1.

La méthodologie proposée est robuste aux conditions d'observation satellitaire qui influent sur la réflectance et la texture des images. Nous avons également montré l'intérêt d'attributs de texture multi-échelles et multi-descripteurs pour la modélisation de la structure forestière.

Fusion de données optique et LiDAR en milieu urbain

En collaboration avec l'équipe MATIS-IGN, nous avons étudié l'apport du LiDAR à retour d'onde complète en complément aux images optiques RVB pour la classification de scènes urbaines. Un ensemble large d'attributs optique et Lidar ont d'abord été calculés d'après ce qui est détaillé dans les sections 3.2.2. Les mesures d'importance des forêts aléatoires ont été ensuite utilisées pour valider l'apport des données multi-sources pour la classification du milieu urbain mais également attester de l'importance de certains attributs pour chacun des classes urbaines (végétation, bâtiments, sol artificiel et sol naturel). Ce travail a été réalisé en 2D en reprojetant les points LiDAR en raster suivant différentes résolutions spatiales [Chehata et al., 2009b] mais également en 3D, en développement des attributs LiDAR 3D [Chehata et al., 2009a, Guo et al., 2011].

Ce travail a été précurseur dans les applications de télédétection et la publication correspondante fait partie désormais des papiers les plus cités du journal IJPRS (International Journal of Photogrammetry and remote sensing). Les publications sont disponibles dans les annexes en section 5.1.2.

Optimisation de la configuration spectrale d'un capteur

Dans le cadre de la thèse d'Arnaud Le-Bris et en collaboration avec l'équipe MATIS de l'IGN et l'ONERA, nous avons développé des méthodes pour la sélection de bandes utiles en imagerie hyperspectrale pour la caractérisation de matériaux urbains. L'objectif était l'optimisation de la configuration spectrale d'un capteur superspectral pour l'étude du milieu urbain en optimisant la position des bandes spectrales mais également leur largeur. Nous avons travaillé sur des méthodes de sélections de bandes supervisées.

Dans notre problématique de conception d'un capteur superspectrale, nous recherchons une solution parcimonieuse en sélectionnant un nombre de bandes très limité par rapport à la donnée initiale. Il semble donc plus pertinent d'utiliser des approches ascendantes comme les algorithmes SFS, SA ou SFFS qui ajoutent progressivement des attributs à la sélection, plutôt que des approches descendantes (SBS ou SBFS) qui partent de l'intégralité des attributs et les éliminent progressivement.

Pour les méthodes d'exploration des attributs, nous avons essentiellement travaillé avec les méthodes SFFS (Sequential Floating Forward Search) et les algorithmes génétiques. Les SFFS est une approche "flottante" qui permet de remettre en question les sélections obtenues aux étapes précédentes, afin d'avoir la configuration optimale lors de l'ajout de nouvelles bandes spectrales. La méthode SFFS a permis de fixer le nombre de bandes optimales afin de comparer de façon équitable plusieurs critères de sélection ensuite. Les algorithmes génétiques ont été utilisés afin de pouvoir inclure différentes contraintes. Par ailleurs, leur principe de fonctionnement conduit à la création de populations de solutions intermédiaires qui nous ont permis de caractériser la stabilité des solutions proposées. Différents critères de sélection ont été comparés sur des jeux hyperspectraux de la littérature (Pavie, Indian pines, salinas), sur la campagne aéroportée UMBRA(projet ANR HYEP) mais aussi sur des libraires spectrales de matériaux urbains. Différentes méthodes de sélection se valent en termes de précisions de classification et renvoient les mêmes groupes de bandes spectrales. La meilleure métrique de la famille *filtre* est Jeffries Matusita alors que dans la famille *wrapper*, les meilleurs résultats ont été obtenus avec une nouvelle métrique basée sur les forêts aléatoires et exploitant la mesure d'importance de variables des RF [Le-Bris et al., 2014a,b, Chehata et al., 2014a].

Nous avons également proposé une méthode innovante de regroupement hiérarchique des bandes spectrales afin d'optimiser leur largeur [Le-Bris et al., 2015, 2017]. L'approche proposée [Le-Bris, 2015] consiste à générer une hiérarchie de groupes de bandes adjacentes, qui sont alors fusionnées. Au sein de cette hiérarchie, les bandes sont progressivement rassemblées selon un critère de fusion en groupes de largeur croissante. Chaque niveau de cette hiérarchie correspond donc à une configuration spectrale couvrant l'intégralité du spectre et composée de bandes de largeur variable. Une sélection de bandes fusionnées est ensuite appliquée à chacun des niveaux de cette hiérarchie, permettant ainsi d'explorer différentes configurations spectrales possibles. Deux approches sont proposées pour effectuer cette sélection de bandes, en prenant en compte la hiérarchie. Ainsi, on a proposé d'utiliser cette hiérarchie de bandes fusionnées comme une contrainte pour l'extraction de bandes et pour limiter la combinatoire.

Plusieurs critères d'optimisation pour la fusion hiérarchique des bandes ont été considérés. Ces critères sont supervisés ou non, selon que la notion de classes a été prise en compte ou pas. Nous avons testé la corrélation entre bandes, la séparabilité entre classes, ou encore l'erreur d'approximation du spectre où les bandes adjacentes sont fusionnées de manière à minimiser l'erreur entre le spectre original et le spectre reconstruit par cette fonction constante par morceaux. Ces travaux sont détaillés en annexe 5.1.3

Autres applications

Par ailleurs, toujours en collaboration avec l'équipe MATIS de l'IGN, nous avons exploré le concept d'importance des variables des forêts aléatoires pour diverses thématiques telles que :

— la cartographie des labours où nous avons considéré différents attributs spectraux et texturaux à partir de données optiques THR [Chehata et al., 2013].

— la délimitation du parcellaire agricole où nous avons utilisé des techniques d'apprentissage supervisé pour apprendre les limites parcellaires. De la même façon, plusieurs attributs spectraux et texturaux à partir de données optiques THR ont été considérés. La sélection d'attributs a été lancée sur plusieurs sous-ensembles d'apprentissage (une sorte de bagging) afin de garantir la robustesse des mesures des importances. Nous avons ainsi pu classer des familles d'attributs et non plus des attributs individuels [Chehata et al., 2015].

— l'évaluation de capteurs optiques pour la cartographie de forêts afin d'enrichir la BD Forêt de l'IGN. A la demande des services de production de l'IGN, nous avons évalué l'apport des données RapidEye RE (5 m de résolution spatiale et une bande red-edge supplémentaire) par rapport aux données orthophotos 50cm RVB de l'IGN pour la classification automatique en 3 classes : résineux, conifères et mixtes. Plusieurs indices spectraux et texturaux ont été dérivés des données RE et ortho. Plusieurs méthodes de sélection d'attributs ont également été comparées. On a pu confirmer dans cette étude que les données étaient bien complémentaires, et que les meilleurs attributs exploitaient la richesse spectrale des données RE et la texture des orthophotos RGB à très haute résolution spatiale [Le-Bris et al., 2013].

— En collaboration avec l'UMR ISPA (Bordeaux INP) sur la forêt des Landes, nous nous sommes intéressés à la segmentation des régions endommagées par la tempête. Dans ce cadre, nous avons proposé une méthode de sélection automatique d'attributs qui permet d'optimiser l'étape de segmentation via une procédure originale, inspirée des procédés de calibrage. La méthode permet d'optimiser 1) les attributs en entrée de l'algorithme de segmentation mais aussi 2) le paramétrage de la méthode de segmentation. La première étape est la construction d'une mire contenant des régions appartenant aux différentes classes à identifier. La mire est calculée pour chaque attribut.Ces mires sont ensuite segmentées individuellement (mono-attribut) et en les combinant (multi-attributs, normalisés) avec différents paramétrages de l'algorithme de segmentation. Le critère de sélection correspond à une métrique d'évaluation de la qualité de la segmentation. Ces travaux sont détaillés dans [Chehata et al., 2014b] qui est annexée section 5.3.1.

Contribution à la communauté

Dans cet axe de recherche, nous avons développé dans le cadre du stage de fin d'études de Walid Bouceffa un module externe en C++ basé sur la libraire OTB (Orfeo ToolBox) permettant de faire la sélection d'attributs de type *wrapper*. Nombreux classifieurs ont été pris en charge : SVM, Random Forests, LDA, Maximum de vraisemblance.. Ce module est désormais diffusé librement au

grand public et il est disponible sur ce lien `https://www.orfeo-toolbox.org/external-projects/`.

3.3 Segmentation

Cette section présente les grandes familles de segmentation d'images ainsi que les techniques de segmentation de nuages de points 3D. Les méthodes qui ont été les plus utilisées dans mes travaux seront présentées ainsi que les diverses applications.

Les techniques de segmentation permettent de partitionner l'image ou le nuage de points 3D en régions homogènes suivant un ou plusieurs critères (radiométrie, couleur, texture, profondeur,...) où l'intersection des régions est nulle. La segmentation permet ainsi d'introduire la notion de cohérence spatiale. Cette information spatiale est à la base des **approches *orientées objet*** qui utilisent les régions segmentées en comparaison aux approches par pixel. Dans la littérature on parle de Object-Based Image Analysis (OBIA) [Blaschke, 2010].

Les méthodes orientées objet ont connu un grand essor avec l'avènement des capteurs satellitaires THR dans les années 2000. Ces capteurs fournissant des résolutions submétriques, un objet se retrouve composé de plusieurs pixels. Une approche par région fournit alors une richesse sémantique permettant de détecter, délimiter mais également de classer ces objets. Par ailleurs un objet complexe, peut être composé de plusieurs objets. Segmenter ces objets élémentaires permet aussi de dériver des raisonnements spatiaux entre ces objets afin de détecter des objets plus complexes [Inglada and Michel, 2009, Vanegas et al., 2013].

3.3.1 Détection d'objets cartographiques

L'application la plus directe de la segmentation est la détection d'objets.

Dans la littérature, aussi bien dans la communauté de vision par ordinateur que télédétection, les travaux se regroupent en trois catégories. (1) **Les méthodes ascendantes (*Bottom-Up*)** se fondent sur l'extraction d'information bas-niveau. Les images optiques submétriques fournissent une description détaillée des milieux d'un point de vue spatial et spectral, permettant l'utilisation simple d'algorithmes de classification supervisée [Senaras et al., 2013]. Cependant, sans modélisation sous-jacente des objets d'intérêt, de telles classifications peuvent êtres très bruitées en raison de la forte hétérogénéité et complexité des milieux. Ces méthodes peuvent être basées sur les contours [Arbeláez et al., 2011, Gioi et al., 2010] ou bien, de façon duale basées sur régions [Gould et al., 2009]. (2) **Les méthodes descendantes (*Top-Down*)** se focalisent directement sur la détection d'objets, par reconnaissance de formes et en utilisant des connaissances a priori sur les arrangements spatiaux et les interactions entre objets [Karantzalos and Paragios, 2009, Benedek et al., 2012]. Cependant, elles sont souvent complexes, hautement paramétrées et avec des temps de calcul rédhibitoires pour des grandes zones. (3) Les méthodes mixtes visent à bénéficier des avantages des deux approches. Dans cette famille, on parlera de **segmentation sémantique** [Arbeláez et al., 2012, Larlus and Jurie, 2008, Kluckner et al., 2009, Kumar et al., 2010]. Certains travaux alternent entre reconnaissance ascendante et détection descendante pour optimiser les deux approches [Harzallah et al., 2009, Ladicky et al., 2010, Maire et al., 2011, Levin

and Weiss, 2006]. D'autres approches vont combiner les techniques de classification et segmentation afin de propager l'information locale obtenue par classification à de grandes zones par un raisonnement spatial structuré [Ok, 2013], en utilisant des optimisations par graphes par exemple. Ces approches de segmentation sémantique sont répandues dans la communauté de vision et commencent à être appliquées en télédétection. Par exemple, [Chai et al., 2012, Youssef et al., 2014b] ont fusionné des méthodes de classification et de reconnaissance de formes pour la détection de bâtiments.

En télédétection, on s'intéressera aux objets cartographiques tels que les bâtiments, les toits de bâtiments, les routes, les rond-points en milieu urbain [Ortner et al., 2007, Youssef et al., 2014b]. En milieu forestier, on s'intéressera plutôt à la détection individuelle d'arbres ou la délimitation de peuplements. En milieu agricole, l'intérêt sera porté sur la détection du parcellaire agricole [Trias-Sanz et al., 2008] ou la délimitation de zones de gestion [Frogbrook and Oliver, 2007, Pedroso et al., 2010] ou encore l'identification de parcelles d'une culture spécifique type oliveraies, vignes, vergers [Mallinis et al., 2008, Delenne et al., 2010, López et al., 2011].

3.3.2 Classification orientée objet

La plupart des méthodes de cartographie par classification supervisée utilisent les méthodes de segmentation suivant trois façons principalement. La première consiste à calculer des descripteurs régionaux (calculés sur le voisinage de chaque pixel) avant une classification pixel à pixel. La seconde passe par un processus de segmentation d'images [Gonzales and Woods, 2008] précédant une classification basée région. La troisième utilise les régions segmentées, a posteriori, dans l'étape de régularisation de la classification. [Fauvel et al., 2013] présente des approches de classification spectro-spatiale basées sur différentes techniques de segmentation. Outre la richesse sémantique des régions par rapport aux pixels bruts, l'approche région permet d'atténuer le bruit de granulation (poivre & sel) et les effets d'ombrage des images satellite [Johansen et al., 2010], mais également d'accélérer le processus de classification en réduisant de façon très significative le nombre d'instances à classifier. En outre, elle est plus robuste aux erreurs de recalage des images qui sont d'autant plus importantes que la résolution spatiale est fine. La principale limitation des approches région est la délimitation des frontières des objets, souvent délicate. Cela explique en partie pourquoi les approches région rapportées dans la littérature ne produisent pas toujours des cartographies plus précises que les approches pixel.

Par ailleurs, la classification de l'occupation à grande et large échelle nécessite de traiter de de grands volumes de données sur de grandes étendues. Les méthodes de segmentations permettent de délimiter le support géographique utile suivant l'application. Par exemple, dans le cas de l'occupation du sol en milieu agricole, l'unité agronomique et hydrologique sera la parcelle ou encore dans le cas de la classification de l'occupation du sol à l'échelle nationale, il apparaît intéressant de stratifier spatialement le territoire afin d'adapter le classifieur à chaque unité (par exemple stratification éco-climatique ou en fonction du paysage). Les méthodes de segmentation permettront de délimiter ces unités afin de les classer plus efficacement. La difficulté dans ce contexte sera

de segmenter des unités intéressantes d'un point de vue sémantique mais qui peuvent présenter de grandes variabilités intra-unités en termes de radiométrie et de texture de l'image.

D'autres verrous scientifique se posent pour les différentes techniques de segmentation : Quels attributs faut-il utiliser en entrée d'un algorithme de segmentation ? quel paramétrage ? et comment évaluer une bonne segmentation ?

Des techniques de sélection d'attributs (cf. section 3.2) sont généralement utilisées pour choisir les meilleurs [Trias-Sanz et al., 2008, VanCoillie et al., 2007]. D'autres travaux essaient d'optimiser le paramétrage de la segmentation [Liu et al., 2012]. Ces méthodes d'optimisation des performances de la segmentation impliquent d'avoir des mesures d'évaluation de la qualité d'une segmentation.

3.3.3 Mesures de qualité d'une segmentation

L'évaluation de la qualité d'une segmentation reste difficile, subjectivement interprétable et elle dépend généralement de l'application finale [Zhang et al., 2008]. [Zhang et al., 2008] présente un état de l'art détaillé des méthodes d'évaluation d'une segmentation. Une évaluation objective est basée sur des mesures de dissimilarité entre deux segmentations. L'évaluation est supervisée si elle dépend d'une vérité terrain, généralement segmentée manuellement à partir d'une image de référence. Cette référence manuelle est elle-même subjective et nécessite un temps de saisie élevé. L'évaluation non supervisée ne nécessite pas une image de référence mais évalue de façon objective une image segmentée en mesurant à quel point elle répond aux caractéristiques attendues par un opérateur humain. Cette évaluation se base essentiellement sur les régions et permet une paramétrisation automatique des algorithmes de segmentation. Les critères d'évaluation non supervisée d'une bonne segmentation peuvent être groupées en trois catégories [Zhang et al., 2008] : 1) mesures d'uniformité intra-région 2) mesure de disparité inter-régions3) mesure sémantique de l'objet(forme) : Ces mesures dépendent de l'application et du type d'image.

Toutes ces mesures s'avèrent efficaces pour la comparaison de différentes paramétrisations d'un même algorithme. Elles sont moins dédiées à la comparaison de différents algorithmes de segmentation et encore moins à la comparaison avec une segmentation subjective humaine. Également, la plupart de ces méthodes sont biaisées par rapport à des sous-segmentations de l'image [Zhang et al., 2008, Neubert et al., 2008].

Les méthodes d'évaluation supervisées restent néanmoins les plus courantes en littérature et en particulier pour la télédétection. Elles utilisent une segmentation de référence obtenue manuellement à partir d'une image, de mesures sur le terrain ou encore à partir d'une base de données existante. De même que les approches utilisées pour la segmentation, l'évaluation de la performance peut être basée sur les contours [Liu and Yang, 1994, Prieto and Allen, 2003] ou bien sur les régions [Zhang, 1996, VanCoillie et al., 2008, Liu et al., 2012]. L'intérêt de l'approche contours est de permettre une plus grande variabilité intra-région mais elle est plus sensible aux problèmes de recalage dans le cas des images de télédétection [Devereux et al., 2004] .

[Zhang, 1996] classe les mesures supervisées de performance de segmentation en cinq catégories que l'on va retrouver dans la littérature sous d'autres

termes ; 1) mesures basées sur le nombre de pixels mal segmentés 2) mesures basées sur la position de pixels mal segmentés 3) mesures basées sur le nombre de régions segmentées 4) mesures basées sur les attributs des régions segmentées et 5) autres mesures.

[Neubert et al., 2008] fait une synthèse des algorithmes de segmentation dédiés aux images de télédétection ainsi qu'un récapitulatif des méthodes d'évaluation supervisée de segmentation. On retrouve essentiellement deux types de mesures ; géométriques et arithmétiques [Liu et al., 2012, Ortiz and Oliver, 2006, Devereux et al., 2004, VanCoillie et al., 2008].

3.3.4 Méthodes de segmentation 2D

Dans ce qui suit, je vais présenter rapidement les méthodes de segmentation que j'ai principalement utilisées et développées dans mes travaux de recherche.

Segmentation hiérarchique

Les données optiques à très haute résolution spatiale rajoutent une double complexité à l'analyse d'images pour la caractérisation des milieux. Une complexité du signal (variabilité radiométrique, rapport signal sur bruit, distorsions radiométriques, géométriques) à laquelle se rajoute une complexité sémantique (nombre d'objets dans la scène, diversité des objets (par exemple arbres, arbustes, bâtiments, routes) mais aussi la diversité ou variabilité de ces objets (différentes formes de bâtiments, différentes espèces végétales, etc). Dans ce contexte de télédétection THR, la segmentation d'une image reste un problème complexe et subjectif. En effet, des régions significatives sémantiquement apparaissent généralement à différentes échelles d'analyse, on retrouve par conséquent les analyses hiérarchiques ou multi-échelles dans les segmentations orientées objet [Baatz and Schape, 2000, Benz et al., 2004, Burnett and Blaschke, 2003, Kurtz et al., 2011]. Les méthodes de segmentations hiérarchiques, mutli-échelles sont des approches ascendantes (*Bottom-Up*) fournissant différentes partitions de l'image à différentes échelles qui minimisent une énergie globale.

Au sein de l'équipe MATIS de l'IGN nous travaillons avec l'algorithme de segmentation multi-échelles, décrit dans [Guigues et al., 2006]. Le principe est de rechercher une famille de solutions, pour $\lambda \in \mathbb{R}^+$ qui minimise l'énergie globale E_λ et d'imposer que ces solutions forment une séquence de partitions ordonnées suivant λ, de la plus fine à la plus grossière, modélisant un processus de description multi-échelles.

$$E_\lambda = E_d + \lambda E_r \tag{3.1}$$

Au niveau pratique, l'idée est de construire une hiérarchie de régions d'une image à différentes échelles en augmentant à chaque fois l'échelle globale λ et en fusionnant pour cela les régions adjacentes dont la fusion diminue l'énergie globale. Pour la fusion de deux régions adjacentes R_1 et R_2 en une seule région R, on définit λ_{app} l'échelle d'apparition de la région R, qui est la valeur minimale à partir de laquelle on a :

$$E(R_1 \cup R_2) < E(R_1) + E(R_2) \tag{3.2}$$

On obtient ainsi une hiérarchie où chaque nœud est indexé par son échelle d'apparition λ_{app} et correspond à une région définie par la fusion d'un ensemble de régions de base connectées. Une méthode naturelle pour la construction d'une hiérarchie est alors de partir d'une partition fine de l'image I qui constitue les nœuds de base de la hiérarchie et de réaliser itérativement la fusion de régions qui a la valeur minimale de l'échelle d'apparition.

La hiérarchie H peut être illustrée comme un arbre où l'axe vertical représente les échelles d'apparition λ_{app}. Les nœuds de base sont obtenus par sur-segmentation de l'image, la racine (i.e le sommet) correspond à l'image entière. La partition finale P est obtenue par une coupe de graphe qui minimise une énergie globale confrontant un terme d'attache aux données $E_d(P)$ et un terme de régularisation $E_r(P)$, $E(P, \lambda) = E_d(P) + \lambda E_r(P)$.

Segmentation par minimisation d'énergie

La segmentation par minimisation d'énergie fait partie des méthodes d'optimisation globale pour les problèmes de labellisation de pixels. De nombreuses publications en vision proposent des approches de minimisation d'énergie pour différentes applications telles que la segmentation d'images [Veksler, 2000, Boykov and Jolly, 2001, Felzenszwalb and Huttenlocher, 2004],la stéréoscopie [Roy and Cox, 1998, Boykov et al., 1998, Kolmogorov and Zabih, 2001], la reconstruction d'objets [Snow et al., 2000], la reconnaissance d'objets [Boykov and Huttenlocker, 1999] ou encore la réalité virtuelle [Thirion et al., 2000].

Cette énergie globale se base sur une formulation Markovienne qui suppose que les interactions entre pixels sont locales i.e. limitées à celles entre voisins. La minimisation d'énergie revient à l'estimation du maximum a posteriori d'un champ de Markov [Geman and Geman, 1984]. L'énergie est composée d'un terme d'attache aux données E_d et d'une énergie de régularisation ou du modèle notée E_r. Le terme d'attache aux données pénalise les solutions incohérentes avec les données observées alors que le terme de régularisation assure une cohérence spatiale. On a alors :

$$E(C) = \sum_{p \in \mathcal{I}} E_{\mathrm{d}}(p) + \lambda \sum_{q \in \mathcal{V}(p)} E_{\mathrm{p}}(p, q) \tag{3.3}$$

avec p un pixel de l'image $\mathcal{I}$, $\mathcal{V}$ le voisinage de p, $\lambda > 0$ un paramètre de pondération entre les deux termes d'énergie et (p, q) une paire de pixels voisins.

Dans nos travaux, nous avons essentiellement utilisé une optimisation par coupes de graphes [Boykov et al., 1999] qui renvoie une solution approximative très proche de la solution globale.

Segmentation par relaxation probabiliste

Dans nos travaux de recherche, nous avons utilisé les méthodes de relaxation probabiliste pour la segmentation d'objets. Ces méthodes permettent de propager des informations dans un graphe. Elle peuvent aussi être utilisées pour la régularisation de résultats de classification par exemple [Smeeckaert et al., 2013].

L'idée fondamentale des méthodes de relaxation sur un graphe [Fu and Yan, 1997] est de faire évoluer de façon itérative des caractéristiques de chacun des sommets en fonction des caractéristiques de ses voisins, de manière à

propager l'information dans le graphe, et ce jusqu'à une position d'équilibre. Dans le cas de la relaxation probabiliste, les sommets sont caractérisés entre autres par des états probabilisés et l'on fait évoluer ces probabilités. On a la relation itérative suivante pour la probabilité à l'itération $k+1$ que le sommet i soit dans l'état λ :

$$p_i^{(k+1)}(\lambda) = p_i^{(k)}(\lambda).\frac{1+s_i^{(k)}(\lambda)}{1+\bar{s}_i^{(k)}} \tag{3.4}$$

et

$$s_i^{(k)}(\lambda) = \sum_{j=1}^{n} d_{ij} \sum_{\lambda'=1}^{m} c_{ij}(\lambda,\lambda') p_i^{(k)}(\lambda') \tag{3.5}$$

$$\bar{s}_i^{(k)}(\lambda) = \sum_{\lambda=1}^{m} p_i^{(k)}(\lambda) s_i^{(k)}(\lambda) \tag{3.6}$$

avec

- Un ensemble de m labels : $\{\lambda_i, i = 1, \ldots, m\}$ pour chaque région.
- $c_{ij}(\lambda,\lambda')$ représente la mesure de compatibilité que le sommet i ait le label λ quand le sommet j a le label λ'. Cette compatibilité peut être considérée comme une probabilité conditionnelle qui satisfait la condition suivante :

$$\begin{cases} \forall\ \lambda,\lambda'\ 0 \leqslant c_{i,j}(\lambda,\lambda') \leqslant 1. \\ \sum_\lambda c_{i,j}(\lambda,\lambda') = 1 \end{cases} \tag{3.7}$$

- d_{ij} quantifie l'influence du sommet i sur le sommet j. Elle vérifie $\sum_j d_{ij} = 1$.

Ainsi, lors d'une relaxation probabiliste, trois paramètres sont à régler :

- La fonction d'influence d_{ij}
- Les probabilités conditionnelles c_{ij}
- L'ensemble des probabilités initiales dénotées $p_i^0(\lambda), \lambda = 1, \ldots m$.

Dans le cadre de la segmentation, la relaxation probabiliste peut être appliquée sur un graphe d'adjacence pour définir la probabilité d'appartenance de chaque région à l'objet par rapport au fond par exemple.

Segmentation Meanshift

La procédure d'optimisation *Mean-shift* a été initialement proposée par Fukunaga et Hostetler [Fukunaga and Hostetler, 1975] puis adaptée par Comaniciu et Meer pour le partionnement (ou *clustering*) et la segmentation d'images [Comaniciu et al., 2002]. L'algorithme *Mean-shift* est une méthode non paramétrique d'analyse d'espaces de caractéristiques complexes et multi-modes, et pour la délimitation de partitions de forme quelconque. Cette approche robuste a été utilisée avec succès pour la classification et la segmentation d'images de niveaux de gris et couleur mais a connu un grand essor dans la communauté télédétection à partir de 2009 [Bo et al., 2009, Wang et al., 2006, Chehata et al., 2014b, Lassalle et al., 2015]. En effet, la segmentation *Mean-shift* ne fait intervenir aucune mesure d'homogénéité contrairement à l'algorithme de segmentation par croissance de régions, elle est donc a priori moins problématique sur les images de télédétection qui sont généralement texturées et

contrastées. De plus, la méthode est non paramétrique, elle n'émet aucune hypothèse sur la distribution des données et ne nécessite pas de fixer le nombre de régions (*cluster*) a priori.

L'algorithme *Mean-shift* regroupe les pixels en les associant à un mode local estimé, c'est-à-dire à un pic de densité dans l'histogramme local [Comaniciu et al., 2002]. Cette méthode explore la fonction de densité de probabilité du ou des attributs considérés dans un voisinage local utilisant conjointement l'information spatiale et radiométrique. Les deux seuls paramètres utilisés sont les résolutions spatiale h_s et radiométrique h_r de la segmentation correspondant aux dimensions spatiale et radiométrique du voisinage. La recherche des modes locaux des pixels s'effectue dans l'espace conjoint des attributs et du domaine spatial à n+2 dimensions où n est le nombre d'attributs considérés. Ces modes sont localisés sans estimer la densité globale qui représente une tâche rédhibitoire en termes de temps de calcul. La méthode est basée sur le décalage récursif de fenêtres de dimension $(n+2)$ vers des modes locaux de la densité des pixels. Ces fenêtres sont centrées en des points initialement choisis de façon aléatoire dans l'image. Ces points sont alors décalés itérativement vers la moyenne de leur voisinage local. Une fois que la localisation du mode est déterminée, la région (ou *cluster*) associée à ce mode est délimitée en fonction de la structure locale de l'espace des attributs [Comaniciu et al., 2002]. Dans un espace à attributs multiples, la segmentation passe par une normalisation préalable des attributs utilisés. Les résolutions spatiale h_s et radiométrique h_r peuvent être déduites de la taille des objets recherchés et de leurs propriétés radiométriques.

Des variantes ont été proposées pour adapter automatiquement la largeur du noyau (kernel bandwith) en fonction des données [Comaniciu et al., 2001], ou bien pour optimiser la recherche des points voisins dans les espaces à grandes dimensions [Georgescu et al., 2003]. Une version hiérarchique du *Mean-shift* a été proposée dans [Paris and Durand, 2007].

3.3.5 Méthodes de segmentation de données 3D

Les techniques de segmentation cherchent à partitionner le nuage de points 3D en zones homogènes et connexes. Les critères d'homogénéité diffèrent en fonction de l'objectif final et peuvent être appliqués à tout attribut dérivé des données 3D LiDAR. Le critère peut être relié aux altitudes en cas de détection de strates forestières par exemple, à la planarité en cas de détection des routes ou de pans de toits par exemple, ou encore à la sphéricité dans le cas d'extraction de couronnes d'arbres.

Deux objectifs sont possibles :

— sur-segmenter le nuage de points, c'est-à-dire le partitionner en segments de petite taille afin de simplifier les étapes suivantes de détection de formes, d'objets, ou pour sa sémantisation. Cette étape est souvent privilégiée dans les chaînes de traitement, car les données LiDAR peuvent être volumineuses. Les techniques standard d'analyse de données et de traitement d'images sont alors adoptées : k-moyennes, k-plus proches voisins, croissance de régions, ligne de partage des eaux, *Mean-shift*, en utilisant la géométrie 3D du nuage de points et souvent un critère de planarité/courbure et/ou d'intensité ;

— détecter directement des objets d'intérêt : couronnes d'arbres en milieux

forestiers, pans de toits, voitures, lignes à haute tension, etc. L'échelle d'analyse est la même que celle des objets d'intérêt. Pour une discrimination fine, cette étape peut également être suivie d'une étape de classification.

La segmentation des nuages de points 3D est essentiellement utilisée en milieu urbain pour la détection d'objets urbains à des fins de classification : la production de cartes d'occupation du sol, ou bien de modélisation 3D urbaine, ou encore en foresterie pour la délimitation des couronnes d'arbres (*ITC : Individual Tree Crown*). En milieu forestier, les données LiDAR à faible densité ($\sim$ 5-10 pts/m^2) permettent d'estimer certaines variables forestières, de détecter les différentes strates d'une forêt [Jaskerniak et al., 2010, Ferraz et al., 2012] ou encore de délimiter les couronnes d'arbres [Strîmbu and Strîmbu, 2015].

La segmentation peut être réalisée en 2D, dans le cas d'utilisation d'attributs LiDAR projetés en 2D, de segmentation d'un MNS (modèle numérique de surface) en milieu urbain, ou d'un MNC (modèle numérique de canopée) en milieu forestier. On retrouve alors toutes les méthodes classiques de segmentation d'image : croissance de régions, K-means, Mean-shift [Ferraz et al., 2012], coupes de graphes fondées sur des critères d'homogénéité spectrale intra-régions ou de forts gradients entre frontières de régions [Strîmbu and Strîmbu, 2015].

La plupart de ces méthodes ont été adaptées à la segmentation des nuages de points 3D. Parmi les méthodes de segmentation de nuages de points 3D les plus courantes, on retrouve l'algorithme de croissance de régions.Une hypothèse d'homogénéité de la région 3D recherchée est spécifiée et peut porter sur tous types d'attributs LiDAR (cf. Section 3.2.2). L'inconvénient d'une telle approche paramétrique est qu'elle ne pourra pas s'adapter aux différents types d'objets rencontrés dans une scène complexe.

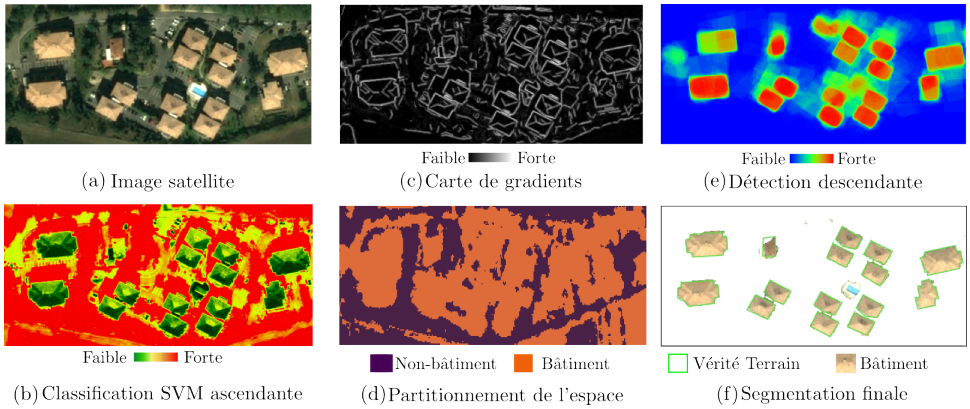
D'autres méthodes de segmentation seront basées sur le partitionnement du nuage de points (clustering) qui correspondent aux techniques de classification non supervisée. En pratique, l'algorithme des K-means est souvent utilisé. Il permet de partitionner automatiquement, les points 3D en groupes homogènes suivant des attributs spécifiques [Chehata and Bretar, 2008]. D'autres approches permettent de réaliser une segmentation non paramétrique des nuages de points 3D telles que l'algorithme du Mean-Shift (sans hypothèse sur la distribution des points 3D dans l'espace des attributs). Il a été appliqué à la segmentation de nuages de points 3D pour la détection des lignes électriques et la segmentation de scènes urbaines [Melzer, 2007], ou encore la détection de strates forestières [Ferraz et al., 2012]. D'autres travaux sur la détection individuelle d'arbres vont utiliser des approches de segmentation par reconnaissance de formes (*template matching*). Le modèle d'arbre sera générique ; ellipsoid [Wolf and Heipke, 2007], conique [Tittmann et al., 2011], blob gaussien [Pirotti, 2010],... D'autres travaux sur des nuages de points 3D terrestres, vont combiner des approches ascendantes et des méthodes de reconnaissance de formes pour avoir une segmentation sémantique de la scène permettant de détecter des objets d'une forme spécifique comme le mobilier urbain (poteaux, panneaux,..), les routes ou encore les véhicules [Golovinskiy et al., 2009, Xiao et al., 2016]. Concernant la validation de ces segmentations 3D, la plupart des travaux utilisent uniquement des métriques arithmétiques comptabilisant le nombre d'objets détectés sans s'intéresser à la délimitation des objets.

3.3.6 Applications

Dans cette section, je vais présenter les différentes applications que j'ai pu traitées en milieu urbain, agricole et forestier et dans lesquelles nous avons proposé des contributions méthodologiques portant sur les techniques de segmentations.

Détection de bâtiments à partir d'imagerie optique THR

Un travail en collaboration avec l'équipe MATIS de l'IGN a porté sur la segmentation sémantique de scène urbaines pour la détection automatique de bâtiments. L'idée est d'améliorer la segmentation par les techniques de reconnaissance de formes. (*cf.* Figure 3.4).



(a) Image satellite (c) Carte de gradients (e) Détection descendante

(b) Classification SVM ascendante (d) Partitionnement de l'espace (f) Segmentation finale

FIGURE 3.4 – *Approche conjointe ascendante et descendante pour la détection de bâtiments.*

Nous combinons 1) une approche ascendante qui exploite les informations bas-niveau de l'image à travers une carte de probabilité de bâtiments obtenue par classification supervisée avec 2) une approche descendante basée sur une modélisation par processus ponctuels marqués (PPM) qui permet de prendre en compte les contraintes géométriques et les interactions spatiales directement au niveau des objets. Une carte de probabilité de bâtiments est alors obtenue par accumulation d'extractions de rectangles à différents paramétrages. La segmentation binaire bâtiment vs. non-bâtiment est réalisée en utilisant une formulation Markovienne combinant les deux cartes de probabilités (ascendante et descendante), optimisée par coupes de graphe [Youssef et al., 2014b,a]. L'article est annexé Section 5.4. Dans ce schéma, nous avons utilisé une fusion d'informations multi-processus ; classification Vs. détection d'objets.

Détection de toits de bâtiments à partir d'imagerie optique THR

Ce travail a été réalisé pendant mes travaux de thèse [Chehata, 2005]. L'objectif était d'extraire des facettes 3D à partir d'un couple d'images stéréoscopiques THR à faible B/H. Deux approches ont été proposées. La première approche (figure 3.5 est basée sur les segments 3D. Les segments 3D sont obtenus par mise en correspondance de segments 2D dans les deux images. Ensuite, l'extraction des plans 3D revient à rechercher des demi-plans pivotant autour des segments 3D extraits maximisant le score de corrélation. Finalement, un algorithme de délimitation des toits sur les plans porteurs permettra d'obtenir les facettes 3D finale. Cet algorithme est basé sur une segmentation par relaxation probabiliste [Chehata et al., 2002].

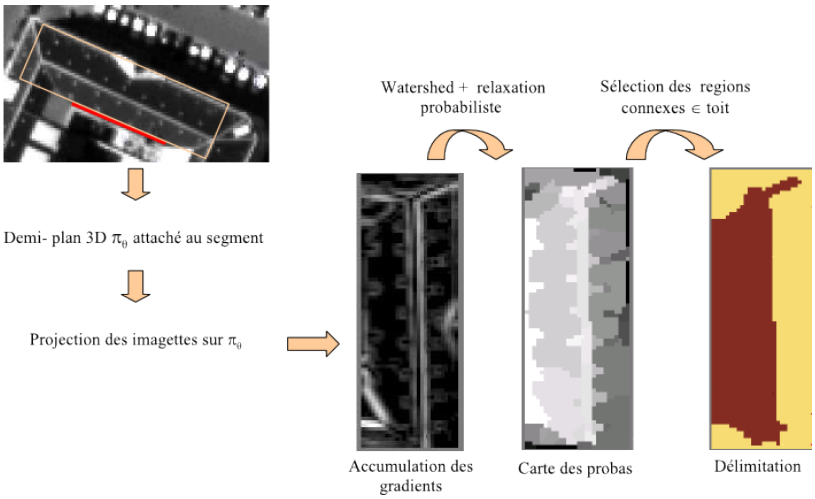


FIGURE 3.5 – *Détection de toits de bâtiments par relaxation probabiliste .*

La deuxième approche est basée sur la segmentation des images en régions. On utilise une description multi-échelles qui fournit plusieurs niveaux d'interprétation de la scène. La deuxième étape est la mise en correspondance des régions où l'on a proposé une approche conjointe ; ascendante pour l'appariement des deux segmentations multi-échelles et descendantes pour avoir la coupe finale correspondant aux régions appariées. La nouveauté, dans notre approche, est l'appariement global des deux segmentations multi-échelles avec une propagation ascendante des contraintes dans la hiérarchie (contraintes épipolaire, d'adjacence), qui rend l'algorithme plus robuste au bruit, aux sur-segmentations des régions et réduit considérablement le temps de calcul.

L'appariement des régions est basé sur des critères géométriques et fournit des associations à différents niveaux de détail. Les deux images sont traitées de façon symétrique. La coupe finale est réalisée par une stratégie descendante (*Coarse To Fine*) pilotée par la mise en correspondance des régions et combine par conséquent des contraintes photométriques (scores de corrélation entre régions) et des contraintes de planéité des régions (surfaces 3D reconstruites à partir des régions appariées). L'appariement global des hiérarchies renvoie deux segmentations non denses des images qui sont appariées entre elles. Les facettes 3D correspondantes sont ensuite reconstruites [Chehata et al., 2003].

L'approche région fournit de bons taux de détection des facettes 3D et surtout un faible taux de sur-détection. L'inconvénient de cette approche est incontestablement le temps de calcul. Par ailleurs le critère de planéité des régions n'est plus fiable sur des petites régions et impose donc une taille minimale de régions finales.

Détection de limites parcellaires

Dans le cadre d'une collaboration avec l'UMR LISAH (IRD, INRA, SupAgro) à travers le projet ANR ALMIRA (Landscape Mosaics of medIterranean Rainfed Agrosystems for a sustainable management of water and soil resources), nous nous sommes intéressés à la caractérisation des éléments anthropiques du paysage à impact hydrologique et tout particulièrement les limites parcellaires. Nous avons choisi une segmentation sémantique du milieu

en utilisant une approche conjointe ascendante et descendante afin de segmenter les parcelles.

Une parcellaire agricole est caractérisée par sa frontière (ligneuse et à fort gradient) et par sa radiométrie que l'on suppose homogène au point de vue couleur et texture. Cette hypothèse ne prend pas en compte les variabilités intra-parcelles.

Nous avons choisi un segmentation hiérarchique construite par des contraintes locales et globales sur les frontières es régions. La coupe finale se fera par approche descendante en utilisant des critères sur la radiométrie des régions. L'énergie globale $E_g$de fusion des régions est formulée comme suit :

$$E_g = \lambda_{gr} E_{gr} + \lambda_{lsd} E_{lsd} + \lambda_{comp} E_{comp} + \lambda_{ori} E_{ori} \tag{3.8}$$

avec E_{gr} l'énergie qui favorise les forts gradients, E_{lsd} l'énergie qui favorise les frontières correspondants à des segments détectés par la méthodes LSD [Gioi et al., 2010], E_{comp} pénalise les frontières complexes et E_{ori} favorise le maintien des limites respectant une orientation globale du bloc de parcelles.

Les premiers résultats sont très encourageants [Youssef, 2013]. Une réflexion porte encore sur les méthodes d'évaluation adaptées au parcellaire agricole en fonction de l'application finale (cf. Section 4.3).

Nous avons comparé ces méthodes avec une segmentation mean-shift hiérarchique [Paris and Durand, 2007]. La figure 3.6 montre le résultat de segmentation d'un parcellaire agricole en utilisant une segmentation mean-shift hiérarchique.

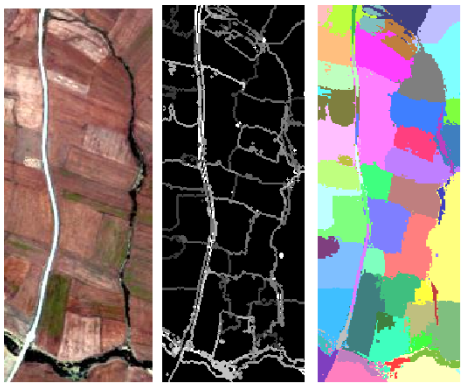


FIGURE 3.6 – *Segmentation de parcellaire agricoles, image Geoeye multispectrale.*

Détection d'infrastructures paysagères à partir de MNS et évaluation de capteurs

Dans le cadre d'une collaboration avec l'UMR LISAH et l'université de Padoue, nous nous sommes intéressées à la détection de terrasses en milieu agricole en évaluant les performances d'un MNS photogrammétriques THR avec un MNS LiDAR aéroporté. Des images stéréoscopiques Pléiades ont été acquises dans le cadre de la RTU Pléiades du CNES. un Modèle Numérique de Surface (MNS) Pléiades été constitué grâce à la suite stéréo-photogrammétrique Mic-Mac par l'IGN, sans régularisation, puis rééchantillonné à $1m$ et nous disposions d'un MNS de même résolution spatiale, acquis par un LiDAR aéroporté (système Toposys) en 2001. Deux méthodes

d'extraction ont été comparées : une approche image par l'algorithme de détection de segments LSD [Gioi et al., 2010] et une méthode géomorphologique basée sur la courbure de surface. Les performances des MNS pléiades sont moins bonnes que celles obtenues avec le MNS LiDAR. Cependant ils permettent de détecter de façon automatique les terrasses dont la pente est supérieure à $2m$ avec un taux de bonne détection de 80%. Par ailleurs, la méthode géomorphologique s'est avérée plus robuste dans le cas de modèles d'élévations bruités.

Dans ce travail, nous avons ainsi évalué les performances de deux capteurs différents optique et Lidar à travers les MNS correspondants pour la détection d'objets cartographiques, plus précisément les infrastructures paysagères type terrasses en milieu cultivé. Les travaux sont détaillés dans [Sofia et al., 2016] et annexés Section 5.2.

Détection de dégâts de tempête en milieu forestier

Dans le cadre d'un projet BQR (Bonus Qualité Recherche) de l' Université Bordeaux 3 et d'une étude Orfeo CNES GT6-Forêt, nous nous sommes intéressés à la cartographie des dégâts de tempête sur la forêt des Landes, la tempête Klaus de 2009. Nous disposions de deux images satellitaires Formosat-2 avant et après tempête. Nous avons proposé une méthode orientée régions pour mesurer le taux de fragmentation des placettes forestières et les classer ensuite en différentes intensités de dégâts. Nous avons utilisé l'algorithme Mean-shift aussi bien pour la segmentation que pour la classification. Nous avons intégré un nouvel indice spatio-temporel dans le processus de détection de changement. Les paramètres de segmentation ont été optimisés par sélection des meilleurs attributs. Les travaux sont détaillés dans [Chehata et al., 2014b] et annexés Section 5.3.1.

Détection de mobilier urbain à partir de données LiDAR Terrestre

Par ailleurs, dans le cadre de l'ANR iTowns, à des fins d'indexation sémantique des nuages de points 3D lidar terrestre acquis par véhicule en milieu urbain, nous avons cherché à segmenter ces données 3D en combinant l'information géométrique du nuage de points 3D et l'information d'intensité des points laser pour détecter le mobilier urbain (poteaux, panneaux) et la végétation (cf. figure 3.7) [Liberge et al., 2010].

Les régions d'intérêts correspondant aux objets verticaux sont extraites par accumulation des points 3D dans le plan horizontal et par l'étude de la distribution 3D locale des points à l'intérieur des cellules de l'accumulateur. Ces régions d'intérêt sont ensuite classées en utilisant une grammaire des formes recherchées. L'objectif est d'apparier un modèle semi-paramétrique de l'objet avec les données laser en utilisant une analyse en composantes principales. classes sont obtenues (poteaux, arbres et autres). La classe poteaux est hiérarchisée en trois catégories (petits poteaux épais, petits poteaux fins et lampadaires). Le taux de bonne classification s'élève à 90 % et le taux de fausses alarmes à 7.5 %.

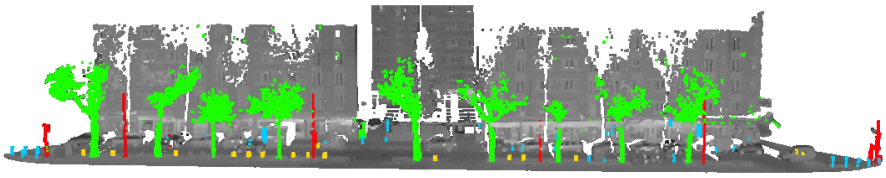


FIGURE 3.7 – *Segmentation de nuages de points 3D terrestre.*

3.4 Classification - Apprentissage automatique

La classification consiste à répartir en fonction d'un critère de similarité un ensemble d'échantillons en plusieurs sous-ensembles appelés classes. Les échantillons à classer sont décrits par des attributs, autrement dit par des vecteurs d'une certaine dimension. On distingue deux types d'algorithmes de classification :

— les algorithmes de classification non supervisée ou clustering, pour lesquels on ne dispose pas de connaissances a priori sur la nature des classes en présence. Il s'agit alors de répartir les échantillons du jeu de données en différentes classes de manière à minimiser la variance intra-classe et maximiser la variance inter-classes de façon à ce que les échantillons appartenant à une même classe soient les plus semblables possibles, et que les échantillons appartenant à deux classes distinctes soient au contraire les plus différents possibles. Parmi les algorithmes de clustering, on distingue les méthodes par agglomération comme la classification ascendante hiérarchique, les méthodes de partitionnement comme les k-means (et leurs différentes variantes : k-médoïdes, nuées dynamiques et ISODATA), et les méthodes paramétriques de type Maximisation de l'espérance pour lesquelles on suppose que les échantillons à classer suivent un modèle de mélange gaussien.

— les algorithmes de classification supervisée ou apprentissage automatique pour lesquels on fournit au classifieur des connaissances sur les différentes classes que l'on souhaite retrouver. Une méthode de classification supervisée comprend donc une phase d'apprentissage ou d'entraînement au cours de laquelle un modèle de classification est calculé à partir d'échantillons labellisés fournis pour chacune des classes. Le modèle calculé à l'issue de cette étape décrit donc la manière dont le classifieur va modéliser les différentes classes. Par la suite, la classification proprement dite consiste à prédire la classe de tout nouvel échantillon en fonction de ce modèle.

Les classifieurs les plus couramment utilisés en télédétection sont les réseaux de neurones [PAI 15], le maximum de vraisemblance, les machines à vecteurs supports (SVM) [Pal and Foody, 2010, Fauvel et al., 2013], ou encore les forêts aléatoires (Random Forests)[Breiman, 2001, Guo et al., 2011]. Dans mes travaux de recherche, j'ai essentiellement travaillé avec les ensembles de classifieurs type forêts aléatoires (RF) et tout particulièrement le concept de marge d'ensemble.

Les classifieurs forêts aléatoires (RF) ont souvent des performances de classification similaires ou supérieures à celles des SVM [Pelletier et al., 2016]. Ils sont rapides en apprentissage et en évaluation. Ils peuvent être facilement parallélisables permettant de traiter de grandes scènes. Par ailleurs, les RF sont multi-classes et ne nécessitent pas de construire plusieurs classifieurs binaires.

Enfin, le classifieur RF est peu sensible au bruit et aux données erronées [Breiman, 2001].

3.4.1 Les forêts aléatoires

Les forêts aléatoires forment une famille de méthodes de classification supervisée, basées sur une combinaison d'arbres de décision. On construit chaque arbre à partir d'une réalisation d'un vecteur aléatoire d'attributs. La décision finale étant prise par vote majoritaire sur l'ensemble des arbres. Une Forêt Aléatoire peut par exemple être construite en générant des sous-ensembles aléatoires d'attributs pour chaque arbre (méthode des sous-espaces aléatoires ou Random Subspaces) et/ou en générant des sous-ensembles aléatoires de données d'apprentissage pour chaque arbre (type bagging). A partir de chaque échantillon, un arbre de décision CART (Classification And Regression Trees) est construit et uniquement un sous ensemble aléatoire de M attributs est utilisé pour la division de chaque nœud. Pour une instance, chaque arbre renvoie un vote unitaire. Le label final est obtenu par vote majoritaire de tous les arbres. Le paramètre M est souvent fixé à $M = \sqrt{N}$ où N est le nombre de variables. L'utilisation d'échantillons, non repris dans le tirage avec remise (bagging) pour la construction d'un arbre particulier, autorise notamment le calcul du taux d'erreurs out-of-bag (OOB). Celui-ci fournit une estimation non biaisée du taux d'erreur en généralisation sans avoir recours à un échantillon test supplémentaire ou à une validation croisée. Les forêts aléatoires sont peu sensibles au bruit et aux données peu informatives, ce qui permet leur utilisation pour la classification avec des vecteurs d'attributs de grande dimension. De plus, à partir des attributs utilisés pour la division de nœuds de chaque arbre, il est possible de déduire une hiérarchisation des attributs et donc de les sélectionner. Les critères d'importance de Gini [Breiman, 2001] et de précision de permutation sont deux mesures distinctes d'importance de variables. Pour l'apprentissage de chaque arbre, à peu près 2/3 des données sont tirées aléatoirement pour construire l'arbre. Le tiers des données restant constitue l'ensemble Out-Of-Bag (OOB). Ces données permettent d'avoir une estimation non biaisée du taux d'erreur en généralisation sans avoir recours à un échantillon test supplémentaire ou à une validation croisée.

3.4.2 Le concept de marge d'ensemble

Marge supervisée

Le concept de marge d'ensemble a été introduit par [Schapire et al., 1998].

La marge d'une instance x est calculée par l'équation 3.9, où v_y est le nombre de votes pour la vraie classe y et v_c le nombre de votes pour toutes les autres classes c. La marge varie entre -1 et +1. Une marge positive ou négative indique que l'instance a été correctement ou incorrectement classée, respectivement. La marge mesure la confiance du classifieur dans sa prédiction. Plus la marge est forte en valeur absolue, plus la confiance du classifieur est importante. Par conséquent la marge d'ensemble est un indicateur de la confiance du classifieur.

$$margin(x, y) = \frac{v_y - \max_{c=1,\ldots,L \cap c \neq y}(v_c)}{\sum_{c=1}^{L}(v_c)} \tag{3.9}$$

Par ailleurs, le signe et la valeur de la marge permettent d'avoir d'autres caractéristiques sur l'instance. Une forte marge positive indique que les arbres sont d'accord sur leurs votes, ce qui correspond à une instance située vers le centre de classe. Inversement, une forte marge négative indique que l'instance a été incorrectement classée avec une forte confiance, ceci correspond probablement à du bruit, des outliers ou encore à des erreurs d'étiquetage. Enfin, une valeur de marge proche de 0 indique que l'instance se trouve sur les frontières de classes.

Dans les problèmes de classification, la plupart des erreurs se produisent sur les frontières de classes. Cependant ces instances contiennent des informations spécifiques sur les classes qui peuvent être intéressantes. Par conséquent, nous avons utilisé une définition non supervisée de la marge d'ensemble permettant également d'évaluer l'ensemble du jeu de données et pas uniquement les données de test.

Marge non supervisée

La marge non supervisée 3.10 correspond tout simplement à la différence de nombre de votes entre la classe votée majoritairement et la seconde classe avec c_1 est la classe majoritaire prédite pour l'instance x, c_2 la seconde classe. La marge varie entre 0 et 1. Plus la marge est faible, plus l'incertitude du classifieur augmente et plus l'instance est proche des frontières de classes et est donc plus informative. Cette marge ne nécessite pas d'avoir le label réel (vérité terrain) de l'instance x.

$$margin(x) = \frac{\max(v_{c_1}) - \max_{c_2=1,\ldots,L \cap c_2 \neq c_1}(v_{c_2})}{\sum_{c=1}^{L}(v_c)} \tag{3.10}$$

Ces concepts de marge supervisée et non supervisée ont été utilisés pour améliorer les méthodes d'apprentissage. En télédétection, plusieurs enjeux scientifiques sur les méthodes d'apprentissage sont liés aux données (*cf.* table 2.2). Dans nos recherches, nous avons essayé de répondre aux points suivants : 1) jeu de données d'apprentissage limité notamment en milieu agricole, où les données sont labellisées par des observations terrain, coûteuses et pas conséquent les jeux de données labellisées sont souvent très limités. 2) jeu de données fortement déséquilibré comme en milieu urbain où la végétation constitue une classe très minoritaire ou bien en agricole avec des types de culture minoritaires. 3) présence de données erronées notamment quand on utilise des bases de données existantes, non mises à jour ou on encore lors du processus de photo-interprétation où des erreurs d'étiquetage peuvent apparaître.

3.4.3 Optimisation de l'ensemble d'apprentissage

Le choix de l'ensemble d'apprentissage est crucial pour les méthodes de classification supervisée. L'optimisation de l'ensemble d'apprentissage doit respecter le principe de parcimonie en limitant la redondance d'information, maintenir une bonne représentativité des classes tout en garantissant de bonnes performances de classification. Différentes études ont montré que la taille et la composition de l'ensemble d'apprentissage conditionnent les précisions de classifications [Foody and Mathur, 2004]. En pratique, généralement des instances pures décrivant chaque classe sont choisies au lieu d'essayer d'avoir des instances qui permettent de mieux discriminer ces classes.

Afin d'optimiser l'ensemble d'apprentissage, nous avons utilisé deux approches différentes suivant des contextes différents. La première approche utilise le concept de marge d'ensemble. Il a été utilisé pour réduire la taille de l'ensemble d'apprentissage ou bien pour gérer des données déséquilibrées. La seconde approche est basée les techniques d'apprentissage actif permettant l'utilisation d'instances non labellisées dans le cas où l'ensemble d'apprentissage initial est très réduit.

Utilisation de la marge d'ensemble

La marge non supervisée traduit la confiance dans le vote du classifieur. Elle peut être utilisée 1) au niveau des données pour améliorer l'apprentissage ou bien gérer des jeux de données déséquilibrés et 2) au niveau des attributs en sélectionnant les meilleurs pour une application visée. La sélection d'attributs a été détaillée dans la Section 3.2. Dans ce qui suit nous nous focalisons sur l'optimisation de l'ensemble d'apprentissage.

Dans le cadre de la thèse de Li Guo sur " Les classifieurs multiples intégrant la marge d'ensemble, Application aux données de télédétection "2011, nous avons travaillé sur l'amélioration des ensembles de classifieurs, notamment les forêts aléatoires (Random Forests) en utilisant le concept de marge.

Généralement, les instances proches du centre de classe sont beaucoup plus nombreuses que celle sur les frontières de classe. Par conséquent, il y a une forte probabilité que les instances centrales soient plus redondantes entre elles. Notre approche sélectionne les instances d'apprentissage les plus informatives i.e les instances de faible marge.

Nous avons alors proposé une méthode appelée "*margin-based bagging*". Un premier ensemble de classifieurs construit par bagging est entraîné. Ensuite, de façon itérative, les instances d'apprentissage de forte marge sont éliminées [Boukir et al., 2013].

Nous avons également proposé une méthodologie de sélection efficace des vecteurs supports d'un classifieur SVM en utilisant la notion de marge d'ensemble [Guo et al., 2010a]. En effet, en considérant l'architecture d'un classifieur SVM, seules les données d'apprentissage proches des frontières de classes sont nécessaires pour construire le classifieur. Par conséquent, supprimer les instances qui sont loin des bords permettrait de réduire considérablement le temps de calcul. Ces instances correspondent à de fortes marges. Autrement dit, les instances de plus faible marge vont correspondre à des vecteurs supports. L'approche a été testée sur différents jeux de données de la base UCI Machine learning. On retrouve à chaque fois autour de 90% des vecteurs supports trouvés par SVM en utilisant juste 25-30% des instances de faibles marges. Les performances sont améliorées surtout pour les classes difficiles.

Techniques d'apprentissage actif

En classification d'images de télédétection, généralement la distribution des classes est supposée stable sur l'image entière. Or en pratique, on est confronté à des différences de distributions entre classes ou l'apparition de nouvelles classes à travers l'image ou bien entre plusieurs images et la construction d'un modèle local s'avère sous optimale pour la classification d'une autre partie de l'image. Dans la littérature, ce problème est intitulé *datashift* et consiste en

un décalage entre les distributions des ensembles d'apprentissage et de test. Ce problème est d'autant plus critique sur des séries temporelles d'images obtenues dans des conditions d'éclairement et d'acquisition différentes.

Par ailleurs, la deuxième contrainte consiste dans le fait qu'en télédétection, la vérité terrain est souvent réalisée par des observations terrain, par conséquent le nombre d'instance labellisées est limité. Dans ce contexte, pour améliorer le classifieur, nous devons exploiter les instances non labellisées.

Les techniques d'apprentissage automatique permettent de lever ce verrou scientifique. Parmi les méthodes, on retrouve l'apprentissage semi-supervisé (apprentissage transductif) [Zhu, 2005, Bruzzone et al., 2006, Maulik and Chakraborty, 2013, Leistner et al., 2009], ou encore actif (ou itératif) [Tuia et al., 2009, 2011, Stumpf et al., 2014]. Dans mes travaux de recherche, j'ai commencé par me focaliser sur les méthodes d'apprentissage actif avec des classifieurs d'ensemble.

L'apprentissage actif consiste à construire, à partir d'un premier échantillon un modèle initial qui est sous-optimal pour de nouvelles données ou images et ensuite de mettre à jour itérativement ce modèle en choisissant de nouveaux échantillons plus informatifs suivant généralement deux critères : l'incertitude et la diversité des échantillons. D'autres travaux ont considéré récemment des contraintes spatiales pour regrouper les échantillons à labelliser sur terrain [Stumpf et al., 2014] ou bien des contraintes en fonction de la topographie et de la facilité d'accès [Demir et al., 2014].

3.4.4 Applications

Dans la partie applications, je vais détailler différents cas d'étude utilisant les techniques d'apprentissage automatique en supervisé et par apprentissage actif , sur des données images mais également 3D.

Classification multi-sources en milieu urbain

En collaboration avec le laboratoire MATIS de l'IGN, nous avons utilisé des données multi-sources (optique et Lidar à retour d'onde complète) pour la caractérisation de milieux urbains. Dans les premiers travaux, nous avons validé avec succès l'intérêt des approches ensemblistes pour la sélection des attributs pertinents en 2D et en 3D et ainsi montré l'intérêt d'attributs LiDAR à retour d'onde complète.

Ensuite, nous avons utilisé le concept de marge pour l'optimisation de l'apprentissage en présence de données déséquilibrées en milieu urbain. Ainsi, nous avons proposé une méthode basée sur la sélection d'instances de faible marge en deux temps. Les faibles marges sont essentiellement localisées sur les frontières de classes et dans les classes minoritaires (cf. figure 3.8).

Une première passe permet de classer les instances de test à forte marge (i.e. généralement les classes majoritaires). Une deuxième passe n'utilisant que les échantillons d'apprentissage de faible marge permet de classer le reste des instances de test. Cette approche permet d'améliorer la précision de classification sur les classes minoritaires tout en maintenant de bonnes performances de classification sur les classes majoritaires. [Guo et al., 2010b].

Ensuite, nous avons généralisé l'approche pour proposer la méthode *"margin-based bagging"*. Sur les données multi-sources urbaines op-

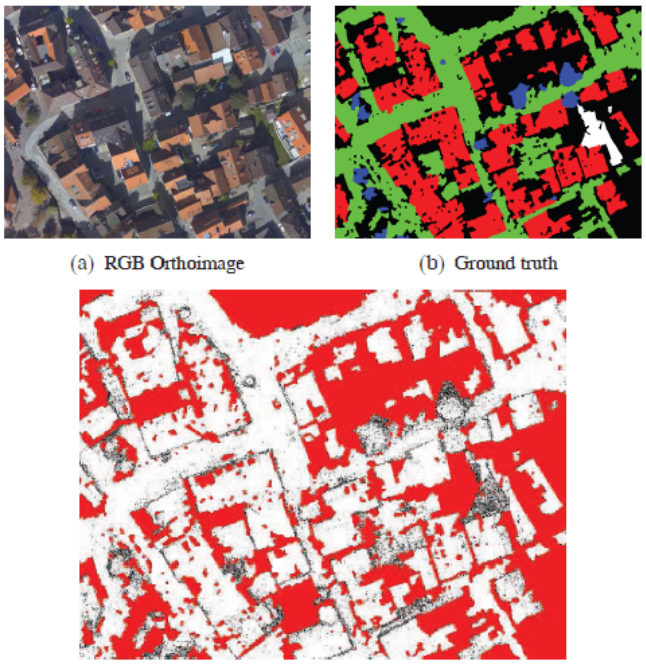


FIGURE 3.8 – *Classification urbaine par forêts aléatoires. Utilisation du concept de marge pour optimiser l'ensemble d'apprentissage dans le cas de données déséquilibrées. Les faibles marges apparaîssent en noir.*

tique/LiDAR, l'ensemble d'apprentissage a été réduit de 15% tout en améliorant la précision globale de classification de 0,6% pour atteindre 94.6 % [Boukir et al., 2013].

Cartographie de l'occupation du sol à grande et large échelle en milieu agricole

Les travaux en cours sont menés en collaboration avec l'UMR LISAH dans le cadre de l'ANR ALMIRA (Landscape Mosaics of medIterranean Rainfed Agrosystems for a sustainable management of water and soil resources) pour la classification de l'ocs à large échelle à partir de données spot. Dans ce cadre, nous avons mis en place une chaîne de traitement automatique de classification supervisée basée sur la librairie OTB prenant en compte une seule image monodate ou bien plusieurs images temporelles intra-saison et dans laquelle nous avons intégré le nouveau module de sélection d'attributs (cf. Section 3.2.6). On s'est rapidement rendu-compte des enjeux spécifiques au milieu agricole : données d'apprentissage réduite, données déséquilibrées, forte variabilité intra-classe et changement des distributions des classes à large échelle.

Nous avons alors proposé la thèse de Ines Ben Slimène dont j'assure actuellement le co-encadrement scientifique et qui porte sur : « Apprentissage actif pour la classification de l'occupation du sol à partir d'images satellitaires à haute résolution spatiale en milieu cultivé ». Sa thèse s'insère dans le cadre d'une cotutelle entre le laboratoire RIADI-ENSI, Tunis et l'UMR LISAH à Montpellier.

Dans ses travaux, nous avons souhaité montrer l'intérêt des techniques d'apprentissage actif (AL) dans un contexte agricole, avec très peu de données d'apprentissage et des contraintes d'observations sur le terrain et en utilisant

des données multispectrales. La nouveauté a été de proposer une approche pratique orientée objet, basée sur la parcelle, unité agronomique, et de prendre en compte les contraintes de terrain. Nous avons montré que l'intérêt des techniques d'AL par rapport à des stratégies aléatoires est surtout confirmé dans le cas d'une précision sémantique élevée (i.e nomenclature ocs niveau 4). Nous avons également proposé une métrique de diversité basée le partionnement par l'algorithme Mean-Shift, qui améliore grandement les résultats en comparaison à des métriques de diversité basées sur les distances tout en maintenant un petit ensemble d'apprentissage (*cf.) Figure 3.9*).

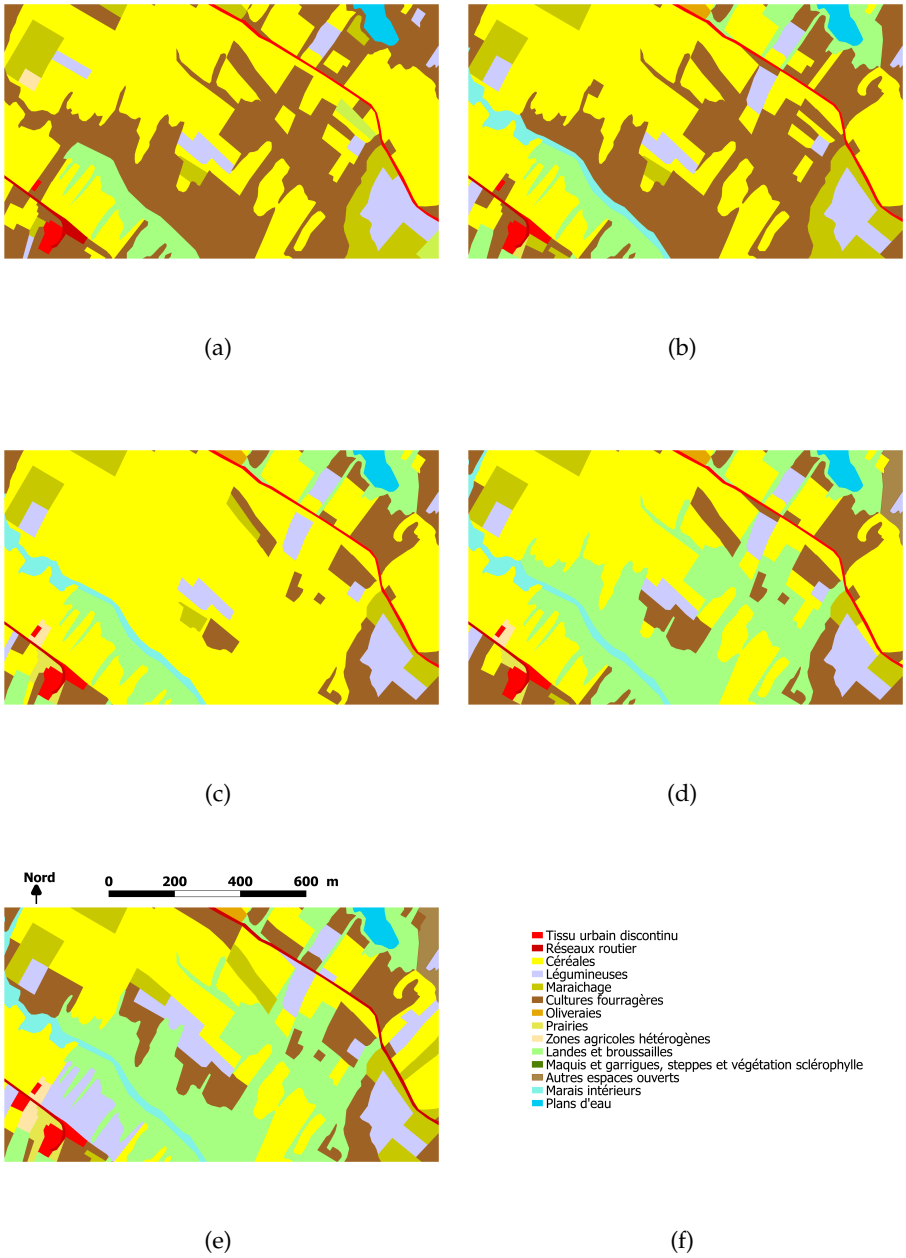


FIGURE 3.9 – *Cartes de pérdictions par AL au niveau L4 (12 ème itération)(a)Stratified Random , (b)Mahalanobis-BT, (e)Mean-shift-Entropy, (f)Mean-shift-BT. (g) Vérité terrain L4.*

Ce travail est en cours de valorisation. Pour la suite de ces travaux, nous

souhaitons lever un autre verrou scientifique qui est la détection de nouvelles classes et l'adaptation de la nomenclature à une cartographie à large échelle.

Délimitation du parcellaire agricole par apprentissage actif des limites

Nous avons développé, dans le cadre du stage de Karim Ghariani une méthode originale de détection de limites parcellaires, à grande et large échelle, basées sur l'apprentissage des limites parcellaires et une classification supervisée binaire "*limite*" Vs "*non-limite*"[Chehata et al., 2015].

En vue d'une délimitation à large échelle, nous avons utilisé les techniques d'apprentissage actif et le concept de marge d'ensemble pour enrichir l'ensemble d'apprentissage par les instances de plus faibles marges afin d'ajuster le modèle du classifieur à d'autres parties de l'image inconnues initialement.

Par ailleurs, une autre nouveauté a consisté à labelliser à chaque itération automatiquement les instances de faible marge en classe "limite" en utilisant une contrainte spatiale par segmentation, i.e une instance de faible marge, correspondant à une frontière de région segmentée, sera labellisée "*limite*" (cf. figure 3.10).

Nous avons testé différentes stratégies d'apprentissage ; apprentissage local (sous-bassin versant), local enrichi par AL, global (échantillonnage sur toute la zone), global enrichi.

On a mis en évidence l'intérêt d'un modèle global si une base de données d'images de limites parcellaires existait. Par contre, si l'on dispose plutôt d'un sous bassin bien renseigné, on privilégiera le modèle local enrichi par AL. Les résultats obtenus constituent des résultats intermédiaires qui peuvent être utilisés dans le cadre d'une segmentation sémantique par minimisation d'énergie. Ainsi, les cartes de marge pourraient être utilisées comme terme d'attache aux données et le terme d'à priori pourrait intégrer des critères de formes sur le parcellaire [Chehata et al., 2015]..

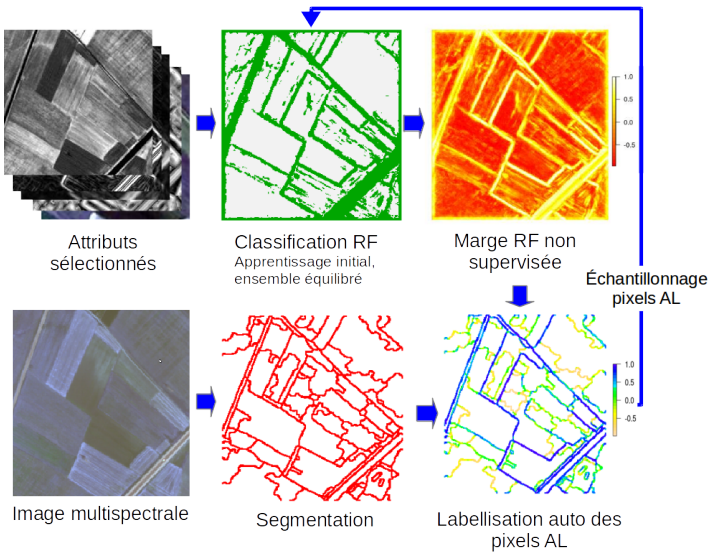


FIGURE 3.10 – *Schéma d'apprentissage actif pour la détection de limites parcellaires à grande échelle.*

Cartographie de la structure forestière à grande échelle : application au pin maritime

Dans la thèse de Benoit Beguet, le concept de marge a également été utilisé pour la correction d'étiquetage dans le cadre d'une cartographie semi-automatique de la structure forestière dans la forêt des Landes. Les cartes de marges permettent de fournir à l'opérateur humain des cartes de confiance du classifieur et donc d'avoir des alertes sur les zones qui sont mal classées malgré une forte confiance du classifieur. Ces alertes sont principalement dues à des erreurs d'étiquetage ou bien à de fortes variabilités intra-parcellaire. Une fois repérée, l'opérateur peut les corriger afin d'améliorer le modèle de son classifieur (cf figure 3.11). Dans notre cas d'étude, nous avons relevé 6% d'erreurs réelles d'étiquetage.

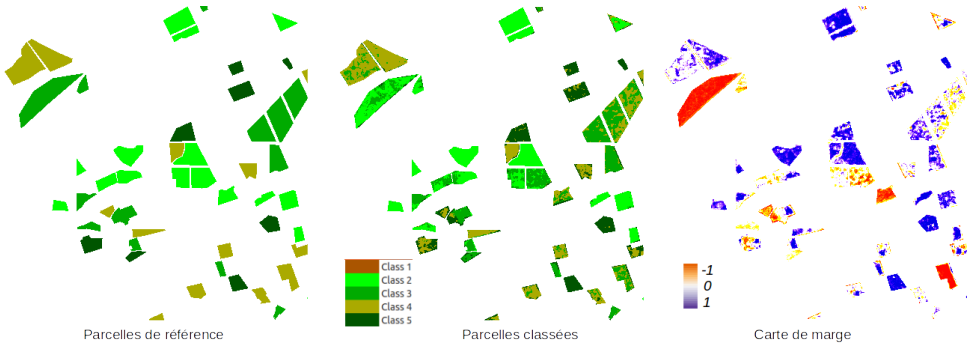


FIGURE 3.11 – *Détection des erreurs d'étiquetage.*

Caractérisation des changements structuraux des couverts de pins maritimes : application à la cartographie des dégâts de la tempête Klaus

Dans le cadre thématique des études Orféo du CNES, nous avons réalisé une étude en collaboration avec l'UMR ISPA de l'INRA Bordeaux sur l'évaluation de l'apport de l'imagerie satellitaire de haute résolution (HR) pour la cartographie des changements des couverts forestiers et particulièrement des dégâts de tempêtes à partir d'images satellitaires HR bi-dates [Chehata et al., 2014b] (cf. figure 3.12).

Cette méthode de détection de changements, complètement automatique et orientée objet, se base sur l'utilisation d'attributs spectraux (indices de végétation), texturaux et temporels qui soient robustes vis à vis de la détection de changements en milieu naturel. Une méthode de segmentation permet de partitionner l'image avant tempête en peuplements de structure homogène et l'image après tempête en zones de changements homogènes. Une méthode d'évaluation quantitative permet de sélectionner le meilleur attribut pour la segmentation . Une méthode de classification multi-temporelle non supervisée permet alors de séparer efficacement les zones endommagées des zones intactes de la forêt grâce à un nouvel indice performant basé sur le taux de fragmentation des régions obtenues [Chehata et al., 2014b]. Nous avons utilisé l'algorithme du Mean Shift [Comaniciu et al., 2002] aussi bien pour la segmentation que pour la classification. La précision de classification obtenue est de 80% par rapport aux données de référence. Les erreurs sont dues principalement à de fausses détections ponctuelles liées aux déplacements d'ombres entre les deux images avant et après tempête, et, à l'omission des zones de dé-

gâts restreintes au sein de jeunes et denses peuplements. Cet article est annexé Section 5.3.1.

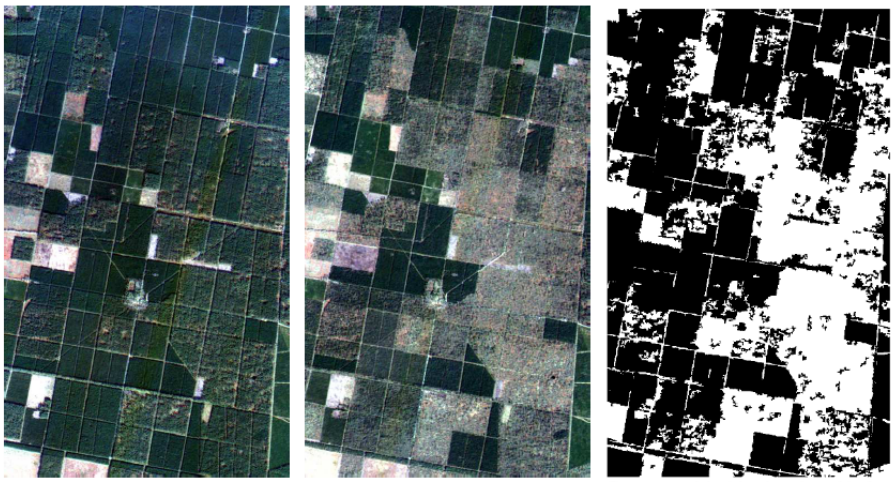


FIGURE 3.12 – *Cartographie binaire des dégâts de la tempête Klaus, à partir d'images formosat2 (8m) bi-dates .*

Classification de nuages de points 3D pour la production de MNT

En collaboration avec l'équipe MATIS de l'IGN, nous avons développé un algorithme hiérarchique de classification de nuage de points 3D, basé sur les K-means, permettant de filtrer les points sol Vs. Sursol. En utilisant les points LiDAR "sol" comme terme d'attache aux données, nous avons proposé une optimisation globale par régularisation Markovienne afin de produire un MNT dense [Chehata et al., 2008, Chehata and Bretar, 2008, Bretar and Chehata, 2010], à une résolution fine de $1m$. Par ailleurs cette approche a été adapté à différents types de milieux ; montagneux en prenant en compte la pente locale et en milieu forestier en adaptant le voisinage des points à la densité du couvert végétal. Pour cela, nous avions combiné les deux sources optique et LiDAR pour proposer un indice inspiré du NDVI et qu'on a appelé HNDVI. Ce travail [Bretar and Chehata, 2010] est annexé Section 5.3.

Classification de nuages de points pour la cartographie du littoral et classification terre/mer.

L'IGN et le SHOM ont lancé depuis 2005 le produit LITTO3D, combinant des données LiDAR topographique et bathymétrique pour la modélisation 3D du littoral Français [Pastol, 2011]. Les données LiDAR sont actuellement utilisées pour classer le littoral en *Terre* Vs. *Eau* permettant de focaliser l'extraction manuelle du trait de côte.

En collaboration avec l'équipe MATIS de l'IGN, nous avons proposé une méthodologie de classification de points LiDAR à large échelle pour la classification terre/mer. La nouveauté a porté sur le développement de nouveaux indices de densité de points permettant de discriminer les points "*Mer*" des points *terre*" le long du trait de côté. Les indices ont été calculés en 2D afin de pouvoir généraliser le processus à large échelle. La deuxième nouveauté a porté sur l'optimisation de l'ensemble d'apprentissage en l'initialisant avec des

traits de côte historiques (i.e. sur des frontières de classes). La régularisation post-classification a été réalisée par relaxation probabiliste en combinant les probabilités d'appartenance aux classes fournies par le classifieur SVM et un a priori contextuel [Smeeckaert et al., 2013].

L'étude a montré la faisabilité d'une approche opérationnelle à échelle régionale et nationale ($\sim$ 567 km^2 répartis sur 9 sites). Des centaines de millions de points LiDAR ont été classés avec une précision de 95% pour la plupart des zones côtières et 90% pour les rivières).

Détection de la desserte forestière à grande et large échelle par données LiDAR

En collaboration avec l'équipe MATIS de l'IGN et dans le cadre du post-doc d'Antonio Ferraz, nous avons développé un algorithme de détection de la desserte forestière en milieu montagneux à partir d'un MNT LiDAR. Ce travail s'inscrivait dans le cadre de projet ANR FORESEE.

La détection de routes est basée sur l'analyse en 2D des points LiDAR à 1m de résolution afin de pouvoir généraliser la méthode à de larges étendues.

Les attributs LiDAR sont basés sur la pente, le gradient de pente et la rugosité du MNT. Des attributs multi-échelles de calcul de pentes permettent de détecter les routes en s'affranchissant d'une détection de leurs contours.

Nous avons utilisé le classifieur des forêts aléatoires et la marge d'ensemble pour générer un masque de focalisation afin d'initialiser les graines du réseau routier et construire un graphe de routes potentielles. Les graines (points) sont initialisées en utilisant des processus ponctuels marqués avec comme terme d'attache aux données la confiance du classifieur (i. marge d'ensemble) et une énergie d'interaction binaire limitant la distance entre graines.

Les arêtes sont ensuite construites en 8-connexité, sur un voisinage d'un rayon de 50m , ce qui permet de créer des arêtes même en cas d'occlusion de la route par la végétation. Les contraintes topographiques (de gradient de pente) et morphologiques (élongation) permettent ensuite d'élaguer les arêtes non réalistes [Ferraz. et al., 2015, Ferraz et al., 2016].

Nous avons encore utilisé une approche combinant classification et détection d'objets. La méthode a été testée à large échelle, sur 1425 km^2 de la forêt des Vosges et validée quantitativement sur 320 km de routes.

Conclusion du chapitre

Dans ce chapitre, j'ai pu détailler les différentes méthodes que j'ai utilisées dans mes travaux de recherche en sélection d'attributs, segmentation et méthodes d'apprentissage automatique. Mes applications ont été principalement conditionnées par mes lieux de rattachement. Étant chercheur invité à l'équipe MATIS de l'IGN depuis 2006, nombreux de mes travaux ont été réalisés en collaboration avec mes collègues du MATIS et le choix des applications a essentiellement porté sur des problématiques IGN. Mes applications en forêt ont commencé en 2008 avec l'Action Scientifique Transverse en Aquitaine et se sont renforcées avec la fusion de l'IGN et de l'IFN en 2012. Mes applications en milieu agricole sont plus récentes et ont commencé en 2012 avec ma mise en délégation à l'UMR LISAH. Ce qui a constitué une sorte de conversion thé-

matique étant donné que le milieu agricole est beaucoup plus complexe avec des enjeux scientifiques qui lui sont propres.

Par ailleurs, mes applications se positionnent dans le nouveau paysage de la télédétection en France. En effet, la communauté de télédétection a connu une restructuration au niveau national grâce au pôle THEIA (pôle Thématique surfaces continentales). THEIA est un réseau d'animation en télédétection qui vise à coordonner les actions de recherche et de formation en Télédétection au niveau régional et national. Ainsi, des CES (Centre d'Expertise Scientifique) produits ont été définis à l'échelle nationale et des ART régionales (Action Régionale Théia) ont été créées dont une en Nouvelle Aquitaine. En parallèle, l'équipex Géosud vise à développer une infrastructure nationale de données satellitaires accessible gratuitement par la communauté scientifique et les acteurs publics en France et dans d'autres territoires.

Ainsi, mes travaux actuels et futurs s'inscrivent dans le cadre de deux CES ; CES OSO (Occupation du Sol) et CES Artificialisation-urbanisation plus récemment (Février 2017). Parallèlement, d'autres travaux récents s'inscrivent également dans l'ART Nouvelle Aquitaine autour de thématiques prioritaires pour la région (la forêt et le littoral).

Perspectives de recherche 4

Sommaire

4.1 Axe apprentissage automatique 94
4.1.1 Optimisation de l'ensemble d'apprentissage 94
4.1.2 Intégration de données temporelles 95
4.1.3 Projets futurs 96
4.2 Axe Fusion de données multi-sources 98
4.2.1 Projets futurs 99
4.3 Axe Segmentation sémantique et méthodes d'évaluation appliquées 101
4.3.1 Projets futurs 102
Conclusion 102

Ce chapitre présente les perspectives de mon projet de recherche afin de répondre aux enjeux précédemment identifiés et lever d'autres verrous scientifiques. Dans ce contexte, les nouveaux projets soumis ou acceptés seront détaillés et positionnés par rapport aux verrous à relever.

Comme on l'a détaillé dans le chapitre 2, les besoins thématiques en cartographie portent actuellement sur la cartographie d'ocs à grande et large échelle avec une meilleure précision sémantique ainsi sur l'analyse plus fine du milieu à travers le calcul d'indicateurs environnementaux dérivés à partir de cette cartographie d'ocs. Différents types d'enjeux scientifiques ont été identifiés et groupés en trois catégories : enjeux liés au contexte opérationnel, enjeux liés aux données et enjeux liées aux méthodes d'apprentissage (Section 2.5).

D'un point de vue méthodologique, on peut considérer que la plupart de ces enjeux relève d'un même verrou commun à lever qui est le **changement d'échelle**. Le changement d'échelle peut être décliné en cinq niveaux. Ces niveaux sont généralement reliés aux caractéristiques techniques des différents capteurs imageurs nécessaires pour chaque type de milieu et chaque besoin thématique :

— **changement d'échelle spatiale** où l'on cherche à produire des cartographies à grande échelle avec des résolutions spatiales de capteurs de plus en plus fines. Les capteurs peuvent être satellites HR ou THR, aéroportés ou encore embarqués sur drones.
— **changement de couverture spatiale** ou d'étendue spatiale qu'on a appelé auparavant la large échelle où l'on cherche à produire des cartographies de l'occupation du sol sur de grandes étendues.
— **changement de résolution spectrale** où l'on doit gérer des données avec des résolutions spectrales de plus en plus en fines, avec des capteurs superspectraux et hyperspectraux.
— **changement de résolution temporelle** avec l'utilisation de séries temporelles avec de grandes fréquences de revisite (1-5 j)
— **changement d'échelle sémantique**, permettant de répondre à un besoin de cartographie fine de l'occupation du sol, plus fine que les nomenclatures des bases de données régionales ou nationales existantes mais également à un besoin d'indicateurs environnementaux plus fins.

Le type de changement d'échelle dépend ainsi des données nécessaires, i.e. du type de capteurs, pour répondre à un besoin thématique précis. Avec l'arrivée des nouveaux capteurs Sentinel S1 (radar) depuis 2014 et S2 (optique) depuis 2015, nous disposons désormais, à l'échelle du territoire français, de séries temporelles d'images gratuites avec une forte résolution temporelle (entre 5 et 10 jours) et une forte résolution spectrale. Ces données optiques sont toutefois sensibles à la couverture nuageuse ce qui peut créer des données manquantes. Et en parallèle nous disposons également d'une couverture nationale monodate à très haute résolution spatiale voire d'images aériennes mais avec une fréquence temporelle beaucoup plus faible pour ces dernières (de l'ordre de 3-4 ans).

Cette abondance de données suscite de nouveaux enjeux liés à l'utilisation de données multi-capteurs, multi-sources et multi-résolutions et par conséquent de nouveaux développements méthodologiques visant à réduire la partie humaine dans les processus de production d'ocs tout en s'adaptant aux différents capteurs et aux différents besoins thématiques. Dans ce contexte, je souhaite, dans la continuité de mes travaux de recherche actuels, approfondir le développement de méthodes de cartographie à grande et large échelle par apprentissage automatique suivant trois axes principaux 1) **l'apprentissage**

automatique en intégrant les données temporelles, 2) **la fusion de données multi-capteurs, multi-sources** et 3) **la segmentation sémantique**.

4.1 Axe apprentissage automatique

Dans la continuité de mes travaux, je souhaite dans cet axe lever des verrous méthodologiques qui sont liés d'une part aux données d'apprentissage et d'autre part à l'intégration de données temporelles, type Sentinel. Dans un contexte appliqué, les données d'apprentissage sont généralement très limitées car réalisées par des observations sur le terrain. Elles peuvent également être déséquilibrées dans le cas d'une application à large échelle ou bien dans des milieux cultivés avec quelques parcelles de types de cultures minoritaires par exemple. Dans le cas où l'on utilise des données temporelles, on est souvent confronté à des données manquantes dans le cas de couverture nuageuse par exemple ou encore à des modifications dans les distributions des classes à cause de conditions d'acquisitions différentes entre différentes dates. Enfin, si le besoin thématique nécessite d'exploiter des bases de données existantes pour constituer les données d'apprentissage ou intégrer des connaissances sur les évolutions temporelles, ces données peuvent contenir des erreurs d'étiquetage du fait de la non mise à jour de ces données ou d'une géométrie imparfaite.

Je décline ainsi les méthodologies de l'apprentissage automatique en deux axes. Un premier axe portant sur l'optimisation de l'ensemble d'apprentissage dans la continuité de la section 3.4.3 et un nouvel axe portant sur l'intégration de données temporelles.

4.1.1 Optimisation de l'ensemble d'apprentissage

Dans le cas d'applications thématiques utilisant des données d'apprentissage limitées ou de présence du phénomène de *datashift* dans les données (*cf.* Section 3.4.3), deux familles d'apprentissage automatique peuvent lever ces verrous : l'apprentissage actif et l'apprentissage semi-supervisé.

Apprentissage actif

Dans la continuité de mes travaux, je souhaite approfondir l'apprentissage actif avec des ensembles de classifieurs en exploitant la marge d'ensemble mais aussi des contraintes spatiales pour répondre aux besoins thématique d'observations sur le terrain. Ces travaux récents Demir et al. [2014], Stumpf et al. [2014], ont été publiés sur l'intégration de contraintes spatiales. Je souhaiterais donc pouvoir optimiser l'ensemble d'apprentissage (en trouvant les plus incertains) en utilisant les ensembles de classifieurs et par des approches globales de minimisation d'énergie impliquant des contraintes spatiales. Ainsi, l'incertitude du classifieur pourrait être utilisée comme attache aux données, la diversité spectrale et les distances entre parcelles utilisées dans le terme de régularisation.

Par ailleurs, les techniques d'apprentissage peuvent être aussi appliquées sur des données temporelles où le modèle du classifieur serait construit à un temps t et appliqué au temps $t+1$ avec enrichissement itératif de l'ensemble

d'apprentissage avec de nouvelles instances de la date $t+1$. Les verrous à lever aussi bien en large échelle qu'en temporel vont porter essentiellement sur les distorsions radiométriques mais aussi sur l'adaptation automatique de la nomenclature avec l'apparition ou la disparition de classes entre les deux dates et la découverte de sous-classes. L'opérateur humain restant dans la boucle de traitement pour labelliser les nouvelles classes découvertes par exemple. Ces techniques seront bien adaptées à la cartographie de l'occupation du sol en milieu cultivé par exemple.

Apprentissage semi-supervisé

Les méthodes semi-supervisées permettent de répondre à un enjeu majeur en télédétection qui est le faible nombre d'échantillons d'apprentissage. En effet, l'apprentissage semi-supervisé vise à exploiter, de façon automatique, conjointement les échantillons étiquetés et non étiquetés disponibles pour concevoir un modèle de classification performant en présence d'un faible nombre de données de référence [Chapelle et al., 2010]. Nous souhaitons dans la suite de nos travaux combiner cette méthode d'apprentissage au paradigme des classifieurs d'ensemble. En effet, l'utilisation des données non étiquetées dans un contexte de classification multiple peut être très intéressante pour contribuer par exemple, en augmentant les ensembles d'apprentissage, à augmenter la diversité des classifieurs de base [Zhang and Zhou, 2013] ce qui augmentera la précision de l'ensemble. .

4.1.2 Intégration de données temporelles

L'exploitation de données temporelles permettra d'une part d'améliorer la précision sémantique de l'ocs et de la végétation en particulier mais également de suivre les milieux afin de détecter des changements abrupts ou continus. Je souhaiterais développer des méthodologies qui permettent de répondre à ces enjeux.

En milieu cultivé par exemple, le besoin thématique sera d'améliorer la détection des types de culture, donc d'avoir une plus grande richesse sémantique de la cartographie de l'ocs et à large échelle.

Un premier apport méthodologique consisterait à améliorer les prédictions d'ocs en prenant en compte des règles d'acteurs construites à partir de données temporelles. Ces règles d'acteurs peuvent être soit temporelles (successions ou rotations de culture), spatiales (type de sol) ou encore contextuelle par le regroupement de culture ou en appliquant le même mode de gestion des parcelles pour un même exploitant. Ces règles peuvent être introduites par un expert agronome à travers des connaissances métier ou bien modélisées automatiquement à travers des séries temporelles d'images ou encore par l'utilisation de bases de données historiques [Osman et al., 2015]. L'utilisation de ces règles d'acteurs tendra à améliorer nettement la cartographie de l'ocs en milieu cultivé par rapport aux attentes des thématiciens.

Un autre besoin pratique en milieu cultivé pourrait être l'utilisation d'une base de données existante d'observations à une date t pour l'appliquer à la cartographie de l'ocs à une date $t+1$. En effet, les données de référence étant obtenues par des campagnes de terrain assez lourdes, la mobilisation des effectifs sur le terrain peut être variable d'une année à l'autre. Par conséquent

l'idée serait d'utiliser les données de référence d'une année pour classer les données de l'année en cours. Pour répondre à ce besoin, je souhaiterais explorer des techniques d'apprentissage par adaptation de domaine pour recaler deux domaines spectraux d'une même scène, correspondant à deux dates différentes. L'idée étant d'adapter le domaine spectral au lieu d'adapter le modèle du classifieur comme c'est le cas en apprentissage actif.

Les données temporelles permettent aussi de suivre les changements abrupts ou continus dans les milieux naturels ou artificialisés. Dans la suite de mes travaux, je souhaite utiliser ces données pour détecter différents types de changement (objets apparus, disparus, ..), hiérarchiser les types de changements mais aussi estimer la surface impactée. La finalité étant de détecter la localisation des changements, de fournir un indice de confiance dans le changement, mais également de proposer des solutions de mise à jour automatique de ces objets modifiés. Ceci permet de disposer d'outils d'aide à la décision pour un opérateur humain par exemple.

D'un point de vue méthodologique, ceci implique de combiner des méthodes de classification multi-temporelle avec des méthodes de segmentation afin de proposer les nouvelles délimitations des objets. Des méthodes basées sur les minimisation d'énergie permettront d'optimiser ce type de problématiques.

4.1.3 Projets futurs

Dans ce qui suit, les prochains projets portant sur l'axe apprentissage automatique seront détaillés.

Cartographie multimodale de l'occupation et de l'usage du sol

En collaboration avec l'équipe MATIS, nous avons proposé un projet ANR MAESTRIA (Multi-modAl Earth obServaTion Image Analysis) pour l'analyse d'images multi-modales qui a pour but de se placer dans un cadre opérationnel, avec une multitude de données ainsi que des bases de données existantes à exploiter pour fournir une ocs à large échelle plus résolue que le produit du CES OSO à $20m$, ç.a.d autour de $2-5m$. Cette cartographie de l'ocs devra être homogène sur tout le territoire français, ne nécessitant pas de données de référence récentes et qui réponde aux besoins thématiques en termes précision sémantique. D'un point de vue méthodologique, ce projet essaiera ainsi de répondre à différentes questions :

- comment fusionner des données temporelles Sentinel avec des données THR multispectrales monodate ?
- comment gérer les grands volumes de données et l'hétérogénéité des capteurs ?
- comment apprendre avec un jeu de données de référence limité et bruité (dans le cas de couvertures nuages, d'ombres ou quand on utilise des bases de données existantes, pas toujours à jour et avec différentes précisions géométriques et sémantiques)
- comment proposer des primitives temporelles robustes ?
- comment gérer les classes sous-représentées ou les nouvelles classes ?

Ce projet se basera sur le produit occupation du sol du CES OSO. D'un point de vue méthodologique, le projet est transversal par rapport aux axes de

recherche que je souhaite. Je me focaliserai sur l'optimisation de l'apprentissage dont la gestion de données déséquilibrées et la fusion de données temporelles et THR. D'un point de vue thématique, on s'intéressera à fournir des services innovants tels que la cartographie de l'usage du sol, la détection de changement, l'adaptation à des milieux complexes tel que le littoral en considérant les besoins des utilisateurs finaux.

Analyse de séries temporelles d'images Sentinel et intégration de connaissances pour la classification en milieu agricole

En collaboration avec l'équipe MATIS de l'IGN, nous avons initié des recherches afin d'exploiter les séries temporelles Sentinel en milieu agricole. Le premier travail consiste à intégrer de façon judicieuse des connaissances a priori afin d'améliorer la classification des différents types de culture. Les connaissances peuvent être de différents types :

— des bases de données existantes type RPG© (Registre Parcellaire Graphique) fournissant le parcellaire agricole et l'occupation du sol de chaque parcelle ou îlot de parcelles.
— des règles agronomiques sur les rotations de cultures et les assolements. Ces règles peuvent provenir implicitement des données RPG.
— des connaissances a priori sur les calendriers culturaux
— des paramètres spatiaux tels que le type de sol, la topographie, les regroupements de cultures

Le verrou méthodologique consiste à intégrer l'ensemble de ces connaissances dans l'analyse des séries temporelles d'images. Pour cela, avec l'arrivée d'un nouveau Chargé de Recherche au MATIS (Loic Landrieu), spécialiste des modèles graphiques, nous proposons d'utiliser les chaînes de Markov à couche cachée et à temps continu qui sont bien adaptées pour assurer cette intégration. Cette collaboration permettra d'aborder d'autres méthodologies jusque là non utilisées dans mes travaux de recherche. Par ailleurs, avec l'arrivée de Sébastien Giordano en tant que chargé de recherche au MATIS, nous nous intéresserons cette fois à l'intégration de données Sentinel 1, radar pour la caractérisation des milieux cultivés. Les données radar n'ont pas été abordés dans mes travaux de recherche jusque là. Le très grand volume de données à traiter impose également la mise en place d'une chaîne algorithmique efficace que l'on puisse passer à large échelle.

Cartographie du littoral Aquitain par classification multiple semi-supervisée

Un nouveau projet TOSCA CLASS-TIP a démarré début 2017 au sein de l'ART Nouvelle Aquitaine sur la "Cartographie du Littoral Aquitain par classification multiple Semi-Supervisée de descripteurs Texturaux extraits d'Images Pléiades". Ce projet est porté par l'EA G&E en collaboration avec l'UMR IMS et la start-up i-sea.

Les attentes de ce projet portent, d'un point de vue méthodologique, sur 1) l'élaboration d'une nouvelle méthode de cartographie de l'occupation du sol par méthode d'ensemble semi-supervisée à partir d'images satellite THR de type Pléiades. 2) l'évaluation du pouvoir discriminant de descripteurs texturaux originaux, issus d'une décomposition en ondelettes, pour la cartographie de l'ocs dans un contexte semi-supervisé. et, d'un point vue applicatif, nous

nous intéresserons à la description fine des faciès du littoral Aquitain, en présence d'un nombre réduit de données étiquetées (transects terrain).

Je me focaliserai dans mon cas sur le premier volet méthodologique et l'application au milieu littoral.

4.2 Axe Fusion de données multi-sources

Afin de répondre aux différents besoins thématiques, nous serons rapidement confrontés en pratique à la disponibilité de plusieurs sources de données (optique/Lidar/MNS voire radar), à plusieurs échelles spatiales (différentes résolutions spatiales) et multi-temporelles. Par conséquent, dans ce deuxième axe de recherche, je souhaite proposer des développements méthodologiques sur les techniques de fusion de données afin d'affiner la cartographie de l'ocs à grande échelle d'un point de vue géométrique et sémantique.

Mes travaux ont essentiellement porté sur la fusion au niveau des attributs en combinant des attributs multi-sources et en étudiant leur importance pour la classification de l'occupation du sol.

Dans mes prochains développements, je souhaite approfondir la fusion de données 1) multi-échelles, 2) multi-sources, et 3) multi-temporelle au niveau des décisions en utilisant des méthodes globales par minimisation d'énergie.

La fusion de données multi-échelles consiste à considérer des capteurs de même type, présentant des résolutions spatiales différentes. Ainsi on peut s'intéresser à la fusion de données riches spectralement (superspectrales ou hyperspectrales) avec des données THR (type Spot ou Pléiades) qui apportent un meilleur niveau de détail géométrique, permettant d'aboutir à une cartographie plus fine sémantiquement et à la meilleure résolution spatiale.

En fonction de la couverture spatiale et du besoin thématique exprimé on peut également s'intéresser à fusionner des données HR ou THR avec des données UHR (ultra haute résolution) acquises par drones. Ce type de fusion prend son sens en milieu cultivé où les objets d'intérêt peuvent être décamétrique voire centimétrique (fentes de retraits, mottes de labour, érosion ravinaire, ...). En effet, une telle fusion prendrait son sens pour valider une cartographie des états de surface avec des mesures de rugosité fines ou encore à prédire des types de sols (argile) avec la présence et la taille de fentes de retrait par exemple ($\sim$10-20 cm).

L'arrivée des capteurs Sentinel a permis d'avoir des données à large étendue, avec une forte résolution temporelle et une résolution superspectrale. Par contre la résolution spatiale reste limitée (10-20m) ce qui en limite l'usage pour certaines applications. Par conséquent, il s'avère intéressant de proposer des méthodes de fusion THR et Sentinel pour améliorer la précision géométrique des ocs dérivées à partir de données Sentinel mais aussi d'améliorer la précision sémantique des produits ocs THR. ce type de fusion sera intéressant en milieu agricole afin d'affiner les limites parcellaires pour les ocs et types de culture mais également en milieu urbain pour proposer des nomenclatures plus fines sur les taches urbaines.

Enfin, pour certaines thématiques, typiquement le milieu forestier ou littoral, il peut être intéressant afin de répondre au besoin de caractérisation du milieu de fusionner des données multi-sources provenant de différents types

de capteurs apportant des informations complémentaires. Ainsi on peut s'intéresser à la fusion de données 2D/3D telle que images optiques vs. LiDAR ou encore images optiques vs. MNS. Le MNS pouvant être obtenu par acquisition LiDAR ou par techniques photogrammétriques. Actuellement, les nouveaux capteurs LiDAR commencent à intégrer cette dimension spectrale avec les LiDAR multispectraux LiDAR XS qui renvoient des mesures d'intensité pour trois longueurs d'onde (vert, PIR et MIR) suivant trois angles de visée 0° Nadir, 3.5^o et 7°, respectivement. La fusion revient dans ce dernier cas au niveau des données.

4.2.1 Projets futurs

Dans ce qui suit, les prochains projets portant sur l'axe fusion de données seront détaillés.

Fusion de données optique multi-capteurs pour la classification en milieu urbain

Dans le cadre du post-doc de Walid Ouerghemmi (projet ANR HYEP), on s'intéresse, pour la classification urbaine, à la fusion de données multispectrales et hyperspectrales afin d'exploiter la haute résolution spatiale et la richesse sémantique respectivement. La fusion est réalisée au niveau décision et est basée sur les probabilités a posteriori d'appartenance aux classes. Différentes méthodes de référence seront testées : fusion bayésienne, logique floue, Dempster-Shafer. De nouvelles méthodes de fusion seront proposées en utilisant le concept de marge d'ensemble.

Ce résultat de fusion sera ensuite utilisé dans un schéma de régularisation globale de la classification afin de gérer les incertitudes sémantique et spatiale. Comme terme d'attache aux données on utilisera la probabilité d'appartenance aux classes de l'image fusionnée et un terme de régularisation privilégiant le même label à des pixels voisins. L'optimisation se faisant par coupes de graphes Hervieu et al. [2016].

Fusion de séries temporelles d'images Sentinel2 et d'image THR pour la classification en milieu urbain

Dans le cadre du pôle Théia, le CES *OSO (Occupation du SOL)* s'occupe de mettre en place une chaîne de traitement nommée IOTA^2 [1] pour la cartographie de l'occupation du sol à large échelle à 20m de résolution. Ce produit est composé de 15-20 classes d'occupation du sol, assez générique pour fournir une cartographie nationale du territoire. Cependant ce produit n'est pas assez précis pour décrire finement l'occupation du sol en milieu urbain ainsi qu'en termes d'usages du sol.

Dans le cadre du CES *Artificialisation - urbanisation* et en collaboration avec le laboratoire LIVE et l'équipe MATIS, nous travaillerons sur la production d'une carte d'occupation du sol en milieu urbain, plus précise que les produits IOTA^2 d'un point de vue sémantique et géométrique. Nous proposons de 1) développer des méthodes de fusion de séries temporelles Sentinel2 avec des données à très haute résolution (THR) type SPOT6-7 ou Pléiades et 2) proposer

1. `http://tully.ups-tlse.fr/jordi/iota2.git`

des indicateurs de l'usage du sol ou des cartes d'usage du sol ou tout au moins une cartographie de l'espace urbain en 5-6 classes.

Pour cela, plusieurs verrous scientifiques sont à lever :

— fusion de données Sentinel 2 et Spot 6. Ce travail s'inscrit dans la suite du post doc de Walid Ouerghemmi.
— Régularisation globale de la cartographie d'ocs en utilisant les résultats de fusion
— la cartographie de l'usage du sol. Deux pistes seront étudiées. Tout d'abord des approches spatiales basées sur les techniques de segmentation ou partitionnement pour développer des indicateurs spatiaux (distance entre objets, densité d'objets, présence de végétation, etc). Dans un deuxième temps, une classification supervisée à partir de bases de données existantes type Corine LandCover ou Urban Allas qui contiennent l'information d'usage du sol mais à moindre résolution spatiale.

Dans ce cadre, la soumission d'un projet TOSCA/ CNES est prévue au premier semestre 2017.

Fusion de données multispectrales et LiDAR pour la détection et caractérisation de strates forestières

Dans le cadre de l'ART THEIA Nouvelle Aquitaine, un projet Région Nouvelle Aquitaine CONFETTI (CaractérisatiON et suivi de l'écosystème ForEstier par Teledetection mulTI-sources et multi-temporelles) a été déposé en Janvier 2017 et qui porte sur la caractérisation fine de la structure de la forêt par analyse de données multi-sources, multi-capteurs, opérant à des résolutions différentes. Ce projet comprend deux volets principaux : 1) un volet de classification multi-temporelle et multi-capteurs (satellites Sentinel 1 et 2) avec des objectifs applicatifs sur la caractérisation de la dynamique forestière et sur la santé des forêts ; 2) un volet de caractérisation de la structure 3D de la végétation de la ripisylve (LiDAR) avec des objectifs applicatifs de modélisation d'un écosystème forestier et de sa biodiversité. C'est sur le deuxième volet que portera ma contribution méthodologique.

Un des facteurs majeurs qui joue un rôle important sur le microclimat du sous-bois forestier mais également sur la dynamique de la biodiversité est la structure géométrique de la végétation. La caractérisation de la structure tridimensionnelle de la végétation, et notamment des différents niveaux de stratification verticale et de la surface occupée par chacune de ces strates, est ainsi essentielle pour ces travaux de recherche. Pour cela, une acquisition LiDAR à forte densité de points (20 pts/ m^2) est prévue. Si ce type d'acquisitions était jusqu'à présent très onéreux du fait de la mobilisation d'avions ou d'hélicoptères, la miniaturisation des capteurs LiDAR permet désormais de les embarquer sur des drones, ce qui rend cette technologie beaucoup plus accessible.

D'un point de vue méthodologique, nous proposons des approches de fusion multi-sources (LiDAR/optique), combinant des informations 2D (LiDAR, photogrammétrie ou réflectance spectrale) avec l'analyse de nuages de points 3D LiDAR pour détecter, caractériser les différentes strates verticales de végétation, et déterminer la variation verticale des densités de surface foliaire. Des métriques de structure 3D seront ensuite développées et serviront à améliorer les modèles existants du fonctionnement de l'écosystème forestier (modélisation du microclimat et de la dynamique de la biodiversité).

4.3 Axe Segmentation sémantique et méthodes d'évaluation appliquées

En utilisant la cartographie de l'occupation du sol à l'échelle nationale, produite par le CES OSO à 20m de résolution, je souhaiterais pouvoir enrichir sémantiquement certaines classes agricoles par l'extraction d'infrastructures paysagères, qu'on appelle les IAE (Infrastructures Agro-Écologiques) et qui peuvent concerner les limites parcellaires mais aussi les haies végétales, les talus , les banquettes, les fossés, etc. La détection des IAE est basée sur la segmentation sémantique de la scène en utilisant des données multi-sources : optiques, LiDAR, photogrammétrique, radar. Le premier verrou sera de combiner des méthodes de classification et de reconnaissance de formes afin de détecter les objets d'intérêts.

Dans la continuité de l'intégration des données temporelles pour la détection de changements surfaciques (par exemple parcellaire agricole), je souhaite exploiter l'information spatiale comprise dans la série temporelle pour délimiter des objets. Dans une formulation énergétique globale, cette information peut être intégrée comme un terme d'attache aux données pour favoriser les contours persistants dans le temps par exemple.

Également une autre piste méthodologique pour améliorer la délimitation d'objets dans le temps, serait d'utiliser les méthodes de segmentation sémantique combinant des résultats de classification (probabilité d'appartenance aux classes) avec des méthodes de segmentation ascendante. Nous avons commencé à utiliser ces méthodes dans le cadre de la détection de bâtiments Youssef [2013] ou de la délimitation de parcellaire agricole. Dans le futur, on pourra intégrer la dimension temporelle.

Par ailleurs, sur l'ensemble de ces méthodes de segmentation proposée, un verrou important portera sur l'évaluation des segmentations en fonction du besoin thématique. En effet, les méthodes d'évaluation de segmentation ont été essentiellement proposées dans la communauté de vision, sur des images d'intérieur ou naturelles. Il s'avère qu'en pratique ces méthodes d'évaluation ne sont pas adaptées à l'évaluation de la précision géométrique des objets cartographiques. En effet, les métriques supervisées qui sont proposées dans la littérature (*cf* Section 3.3.3) sont pour la plupart soit arithmétiques soit géométriques (calcul de recouvrement). Ces méthodes se concentrent sur la détection des objets, leur localisation et non pas sur leur précision de délimitation ou leurs relations topologiques.

Lors de nos travaux sur la délimitation de parcellaire agricole, nous nous sommes alors posé cette question : comment évaluer le parcellaire obtenu, de façon à ce qu'il réponde au besoin thématique exprimé qui était leur utilisation dans un modèle hydrologique ?

Nous avons conclu que l'échelle d'analyse ainsi que les méthodes d'évaluation dépendait de l'application finale. Ainsi, l'échelle globale (carte de densité, carte de chaleur, ..) permet de valider des arrangements spatiaux, des motifs répétitifs dans le paysage. L'échelle locale quant à elle peut être analysée par deux points de vue ; 1) par les contours si nous nous intéressons à la délimitation du parcellaire, à la topologie des arrangements par exemple dans le cas d'utilisation dans les modèles hydrologiques,.. 2) par les surfaces si l'on s'intéresse aux déclarations de surfaces cultivées,..

Je souhaiterai donc approfondir cette problématique d'évaluation de seg-

mentations qui reste difficile pour répondre à un besoin thématique donné, notamment en milieu agricole pour l'évaluation du parcellaire agricole.

4.3.1 Projets futurs

On retrouve l'axe segmentation sémantique dans les deux premiers axes. En effet, elle sera développée dans l'axe apprentissage avec l'intégration de séries temporelles et l'exploitation de l'information spatiale dans ces séries. Elle sera également utilisée pour la détection d'objets. Nous souhaitons en collaboration avec l'équipe MATIS, et dans la continuité de nos travaux communs (stages de Karim Ghariani et Mahmoud Sidi Youssef) Chehata et al. [2015], Youssef et al. [2014a] combiner des méthodes de classification supervisée et des méthodes de segmentation dans un cadre de régularisation globale incluant aussi la régularisation temporelle afin de détecter les objets urbains (CES urbanisation) ou encore le parcellaire agricole. On intégrera également la segmentation sémantique dans la fusion de données 2D/3D par exemple pour la détection de couronnes d'arbres ou de strates forestières. Par ailleurs, sur les aspects d'évaluation de segmentation, nous avons proposé avec l'UMR LISAH un stage PFE de Meriem Mejri pour l'évaluation du parcellaire agricole où l'on a proposé une version localisée de l'indice Rand Index. Au niveau des projets, nous avons déposé un projet TOSCA TRIPL (Télédétection hautement Résolue d'Infrastructures Paysagères Linéaires). Ce projet n'a pas été retenu car il était mal positionné par rapport à la dynamique nationale en 2014. Il devra par conséquent, dans le futur, être mieux articulé par rapport au CES OSO.

Conclusion du chapitre

Dans ce chapitre perspectives, nous avons tracé les grandes lignes des futurs travaux de recherche. Ils restent dans la continuité de mes travaux actuels. La nouveauté sera l'intégration de la dimension temporelle avec l'arrivée des Données Sentinel2 et les besoins thématiques correspondants, en utilisant les méthodologies développées dans mes travaux mais aussi de nouvelles méthodologies toujours basées sur les graphes en collaboration avec de nouveaux chercheurs de l'équipe MATIS de l'IGN. Enfin, les actions à court et moyen terme démarrées et en cours de montage ont été présentées ainsi que leurs verrous scientifiques et des pistes méthodologiques ont été proposées afin de les lever.

Annexes 5

5.1 Axe Sélection d'attributs

5.1.1 Sélection d'attributs pour la modélisation de la structure forestière

ISPRS Journal of Photogrammetry and Remote Sensing 96 (2014) 164–178

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Automated retrieval of forest structure variables based on multi-scale texture analysis of VHR satellite imagery

Benoit Beguet [a,b,*], Dominique Guyon [b], Samia Boukir [a], Nesrine Chehata [a]

[a] EA 4592 G&E, ENSEGID, University of Bordeaux, 1 Allée F. Daguin, 33607 Pessac Cedex, France
[b] INRA, UMR1391 ISPA, 33140 Villenave d'Ornon, France

ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 28 April 2014
Accepted 8 July 2014

Keywords:
Forestry
Multiple regression
Feature selection
Texture
Multi-scale
Multi-resolution
Pléiades
Quickbird

ABSTRACT

The main goal of this study is to design a method to describe the structure of forest stands from Very High Resolution satellite imagery, relying on some typical variables such as crown diameter, tree height, trunk diameter, tree density and tree spacing. The emphasis is placed on the automatization of the process of identification of the most relevant image features for the forest structure retrieval task, exploiting both spectral and spatial information. Our approach is based on linear regressions between the forest structure variables to be estimated and various spectral and Haralick's texture features. The main drawback of this well-known texture representation is the underlying parameters which are extremely difficult to set due to the spatial complexity of the forest structure.

To tackle this major issue, an automated feature selection process is proposed which is based on statistical modeling, exploring a wide range of parameter values. It provides texture measures of diverse spatial parameters hence implicitly inducing a multi-scale texture analysis. A new feature selection technique, we called Random PRiF, is proposed. It relies on random sampling in feature space, carefully addresses the multicollinearity issue in multiple-linear regression while ensuring accurate prediction of forest variables. Our automated forest variable estimation scheme was tested on Quickbird and Pléiades panchromatic and multispectral images, acquired at different periods on the maritime pine stands of two sites in South-Western France. It outperforms two well-established variable subset selection techniques. It has been successfully applied to identify the best texture features in modeling the five considered forest structure variables. The RMSE of all predicted forest variables is improved by combining multispectral and panchromatic texture features, with various parameterizations, highlighting the potential of a multi-resolution approach for retrieving forest structure variables from VHR satellite images. Thus an average prediction error of ~1.1 m is expected on crown diameter, ~0.9 m on tree spacing, ~3 m on height and ~0.06 m on diameter at breast height.
© 2014 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Many studies have focused on estimating forest parameters from optical remote sensing data since the early years of satellite imagery. The methods consist usually in extracting one or several predictive variables for estimating tree or stand level variables that are useful for forest resources inventory and management such as tree height, stem diameter, density, basal area, stem volume or biomass (Kayitakire et al., 2006; Feng et al., 2010; Proisy et al., 2007; Ozdemir and Karnieli, 2011; Song et al., 2010; Beguet et al., 2012). These studies are generally based on linear single regression modeling and only few multiple regression modeling. The accuracy in retrieving forest stand variables depends on the image spatial resolution. Hyyppa et al. (2000) observed that medium or high resolution multispectral satellite imagery such as SPOT-4 (20 m resolution) or Landsat TM (30 m resolution) leads to a lower correlation performance, in comparison with sub-metric aerial photography. Using 5 m and 10 m SPOT-5 images, Wunderle et al. (2007), Wolter et al. (2009), Castillo et al. (2010) retrieved some forest stand attributes (such as crown diameter), exploiting image texture, with a good accuracy.

Over the last decade, a growing number of Very High Resolution (VHR) multispectral satellite images from various sensors has

* Corresponding author at: EA 4592 G&E, ENSEGID, University of Bordeaux, 1 Allée F. Daguin, 33607 Pessac Cedex, France. Tel.: +33 5 57 12 10 36; fax: +33 5 57 12 10 01.
E-mail address: benoit.beguet@ipb.fr (B. Beguet).

http://dx.doi.org/10.1016/j.isprsjprs.2014.07.008
0924-2716/© 2014 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

become available such as Pléiades, Quickbird, Geoeye, WorldView or Ikonos which provide sub-metric spatial resolution (0.5–1 m in panchromatic band, 2–4 m in multispectral bands).

VHR imagery provides a meaningful textural information. Various texture representations were proposed in literature and applied to the remote sensing data for a wide range of applications such as urban mapping (Soe and Tyler, 2004; Dell'Acqua and Gamba, 2003; Pacifici et al., 2009), image segmentation (Trias-Sanz et al., 2008; Gaetano et al., 2009) or vegetation structure mapping and habitat use (Wood et al., 2012; Tuttle et al., 2006).

Some recent studies (Kayitakire et al., 2006; Feng et al., 2010; Proisy et al., 2007; Ozdemir and Karnieli, 2011; Song et al., 2010; Beguet et al., 2012; Beguet et al., 2013; Gomez et al., 2012; Tuominen and Pekkarinen, 2005) have shown the potential of VHR imagery for forest inventory applications thanks to the strong relationship between forest spatial structure and image texture at stand level. Texture analysis of VHR satellite images generally applied to forest inventory can be divided into three main approaches (Maillard, 2003): wavelet or Fourier-based, variogram-based, and GLCM-based (Grey level Co-occurrence matrix) methods. Couteron et al. (2005), Proisy et al. (2007), Barbier et al. (2010) used a 2D Fourier periodogram to model canopy texture. Ruiz et al. (2004), Regniers et al. (2013), Van Coillie et al. (2007) used wavelets for forest structure analysis. A variogram is useful to explore the relationship between image texture and forest structure (Guyon and Riom, 1996; St-Onge and Cavayas, 1997; Franklin et al., 2001; Wulder et al., 1998; Song et al., 2010). Its main drawback is the requirement of a good non-linear model to fit the observed variogram before extracting texture indicators such as sill, range and nugget. The second order statistics derived from Grey Level Coocurrence Matrix (GLCM) as defined by Haralick et al. (1973) are the most used texture features in forestry remote sensing literature (Franklin et al., 2001; Chehata et al., 2011; Boukir et al., 2013) and provide good performances for estimating forest parameters (Kayitakire et al., 2006; Wunderle et al., 2007; Wolter et al., 2009; Castillo et al., 2010; Ozdemir and Karnieli, 2011; Beguet et al., 2012). Generally, these parameters are set to fixed and hence unoptimal values. To tackle this major issue, an automated feature selection process that explores a wide range of texture parameter values is a worthwhile solution to investigate which, to the best of our knowledge, has not been proposed yet in literature. This automated spatial parameter tuning has an appealing property : it provides texture measures of diverse spatial parameters hence implicitly inducing a multi-scale texture analysis. It is well known that texture analysis is generally more relevant in a multi-scale context (Gonzales and Woods, 2008).

Moreover the existing studies generally involve only one spatial resolution, considering exclusively panchromatic or multispectral data. To our knowledge, only Wolter et al. (2009) have combined texture features from both kinds of data with different spatial resolutions, but via variograms and not GLCM.

Besides, the impact of image acquisition conditions on the forest structure retrieval from texture information is a critical issue. In fact, both view and sun angles influence the image texture, due to the interaction of radiation with vegetation structure. They particularly determine the length and the orientation of the shadows of crowns cast on ground, the fraction of sunny or shadowed crowns, and also the apparent radius or length of crown viewed by the sensor. Barbier et al. (2010) showed that the bidirectional variation in texture is to be accounted for when using a number of images with various angle configurations. In addition, image texture can significantly vary across the seasons due to phenological changes in vegetation structure, thus indicating that care is to be taken when measuring texture at different phenological stages (Culbert, 2009).

In this paper, we aim to fully exploit the potential of texture features extracted from VHR satellite images using a GLCM-based approach for estimating some typical forest stand variables, with a particular emphasis on automated parameter tuning, one of the major issues in texture analysis. The study focuses on the retrieval of the following forest structure variables: crown diameter, tree height and tree density or spacing which lead a spatial resolution-dependent image texture. The stem diameter at breast height is also considered, since it is easily measured on the field and it indirectly contributes to image texture via its correlation with the crown structure variables such as crown diameter. The top-tree height also contributes to image texture since it impacts on the length of crown shadows according to the solar elevation or the apparent crown length viewed by the sensor according to the view zenith angle.

The main objective of our work is to design and test an automated process for the selection of GLCM texture features and their optimal parameterization that is able to predict the forest variables with a satisfactory accuracy regardless of the season or the angle configuration of the remote sensing data. To achieve optimal results, numerous combinations of panchromatic and/or multispectral features were tested through multi-variable linear regressions involving field-measured forest variables. As collinearity is a very perturbing problem in multi-linear regression, this issue is carefully addressed through an innovative variable subset selection algorithm we called Random PRiF. A comparative analysis is carried out with other well-established variable subset selection techniques used in multi-variable regression, a classic dimensionality reduction: PCA (Principal Component Analysis) (Manly, 1994) and a recent one that is more adapted to prediction: LARS (Least Angle Regression) stepwise.

The developed automated forest variable retrieval scheme was assessed on stands of maritime pine (*Pinus pinaster* Ait.) of the forest of "Landes de Gascogne" in south-western France. An accurate estimation of the forest variables is aimed. Indeed, covering about one million hectares (75% of the total area of the region), these maritime pine stands are economically important at regional and national levels and thus an accurate inventory of wood resource is required. Two sites covering a large part of the structure diversity of this forest were considered, using VHR Quickbird and Pléiades images.

2. Material

2.1. Study sites

Both study sites are located in south-western France within the largest European maritime pine (*Pinus pinaster* Ait.) forest (Fig. 1), so-called *the forest of Landes de Gascogne*, which covers approximately one million hectares in a nearly flat area except for the coastal dunes. This forest consists of maritime pine even-aged stands which are intensively managed. The pine trees are planted in rows usually 4 m apart. The regeneration techniques have changed in recent years. Now the trees are generally regularly planted along the rows with a low density (~1000 trees/ha, i.e. within-row spacing ~2.5 m). Previously, the trees were always sowed with a higher density in the row (~5000 trees/ha, i.e. within-row spacing ~0.5 m). The trees are periodically thinned after clearing the understory vegetation. The clear-cutting occurs mostly when the pine trees are 50 years old.

The Nezer site, whose area is around 60 km^2, is managed quite uniformly since it covers only two large tree farms. The stand size is large, the mean area is approximately 12 ha with a maximum of 50 ha. The stands are mostly rectangular and often circled by fire

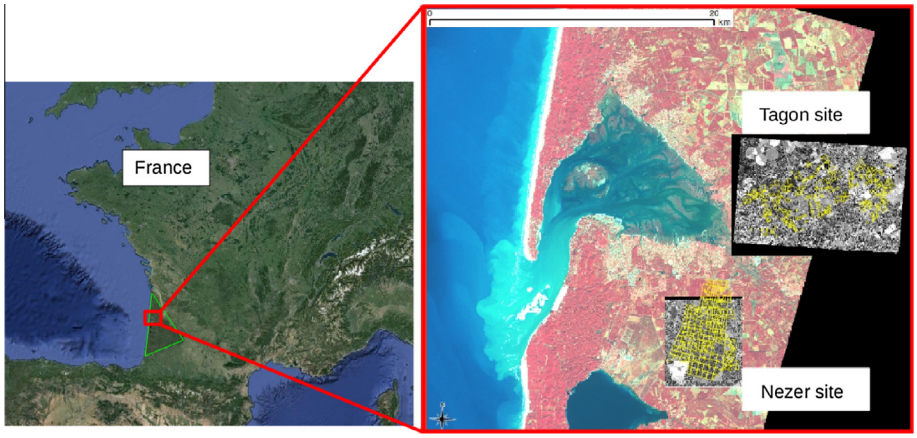


Fig. 1. Location of Nezer and Tagon study sites.

lanes or roads. The pine trees rows are oriented West–East in most of stands.

The Tagon site, which covers about 80 km^2, is more complex and more heterogeneous than the Nezer one, since it is managed by many forest owners. The stands are smaller (mean area ~7 ha, maximum ~40 ha) and displays a larger variability in geometry, row orientation, and forest structure.

Both sites are highly representative of the whole maritime pine forest diversity in terms of forest structure, understory species composition, silvicultural and management practices.

2.2. Field data

For both sites, four forest structure variables were measured on field: diameter at breast height of trunk (*Dbh*), height of the tree top (*Ht*), crown diameter (*Cd*) and stand density (*Nah*). *Sp* is a tree spacing index that was calculated from *Nah* using a non linear function (*cf.* Eq. (1)) assuming each tree is located on the apex of an equilateral triangle. The field measurement campaigns have been driven by GIS using aerial photographs and the forest management data which were available. Thus several forest structure classes have been designed considering tree spatial organization and dimensions, in order to sample uniformly the whole range of forest variables variation. No measurements were performed on the very young stands (<4 year), where the tree cover fraction is very low and the undergrowth vegetation is often taller than the pine trees.

$$Sp(m) = \sqrt{\frac{20,000}{\sqrt{3} * Nah(\text{tree/ha})}} \quad (1)$$

For Nezer site, the measurements were made on 12 stands at the end of the year 2003. The trees sampled in each stand were located within a 80 m × 80 m square plot representative of the forest structure in the 200 m × 200 m area that encloses it.

For Tagon site, the measurements were made on 111 stands during summer 2012 at the end of the pine tree growth period. The trees sampled in each stand were included in a 20 m×20 m square plot representative of the forest structure in the 80 m × 80 m area that encloses it.

The variation range of the forest variables (mean values per stand for *Cd*, *Ht*, *Dbh*) are summarized, for each site, in Table 1 and Fig. 2. Fig. 2 highlights the allometric relationships between variables *Cd*, *Sp* and *Ht*, i.e. their intrinsic correlations to the species and their correlation related to tree growth, site fertility and past silvicultural operations.

Table 1
Variation range of forest variables (Nezer: 12 sampled stands, Tagon: 111 sampled stands). Mean values per stand are given.

	Cd(m)	Ht(m)	Nah(tree/ha)	Sp(m)	Dbh(m)	Age(year)
Nezer site						
Min	2.95	9.3	189	3.04	0.15	13
Max	7.81	24.7	1253	7.81	0.46	51
Tagon site						
Min	0.77	1.7	150	1.31	0.02	4
Max	10.69	26.2	6729	8.77	0.56	68

2.3. Remote sensing data

VHR satellite images from two different sensors (Quickbird and Pléiades) were used. These images have a similar spatial resolution in similar spectral bands: ~0.5–0.6 m in panchromatic (Pan), and ~2–2.5 m in multispectral mode (MS). On Nezer site, one Quickbird image was processed and analysed. It was acquired on the 6th of October 2003. On Tagon site, two Pléiades images were acquired respectively on the 26th of June 2012 and the 8th of August 2012. All the images were geo-referenced and ortho-rectified. Their characteristics are given in Table 2.

This image set gives some repeated observations in different conditions of phenological stage of vegetation (summer/autumn), solar elevation (low/high) and angles of view (oblique/very oblique). The differences are higher between Nezer and Tagon datasets. For the Quickbird image acquired on Nezer in autumn, the understory vegetation was senescent with a low green leaf area whatever its species composition. Its contribution to the image reflectance spatial variation and consequently to image texture was probably low. On the contrary, the Pléiades images were acquired on Tagon in summer season when the green leaf area of the undergrowth vegetation is maximal. However, its impact on the reflectance can strongly vary in space according to the species (fern, heather, grass such as molinia and other) and their spectral properties.

3. Methods

The global framework of forest variable retrieval methodology is presented in Fig. 3. Test frames that are composed of image samples (reference stands) are constructed from both panchromatic and multispectral bands. Texture features with various parameterizations are then calculated on test frames. Median values are kept per image sample. These test frames are used for feature selection based on linear regression to keep the best texture features that are

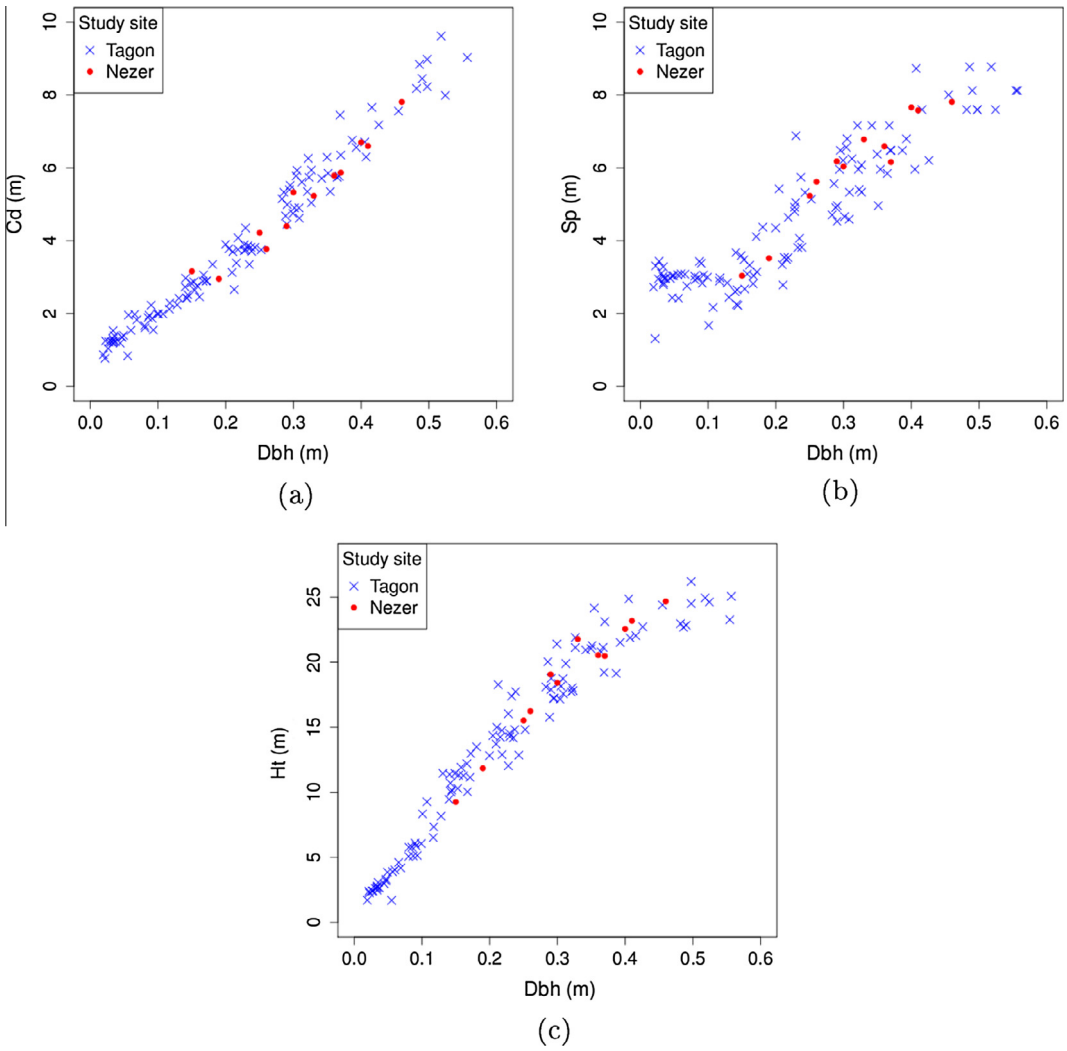


Fig. 2. Field measured forest variables on both sites: crown diameter (Cd), tree height (Ht) and tree spacing (Sp) as a function of diameter at breast height (Dbh).

Table 2
Image acquisition parameters.

	QuickBird 06/10/2003	Pléiades 26/06/2012	Pléiades 08/08/2012
PAN-MS resolution (m)	0.62–2.48	0.5–2	0.5–2
Encoding (bits)	11	8	8
Panchromatic	450–900 nm	450–900 nm	450–900 nm
Blue	450–520 nm	430–550 nm	430–550 nm
Green	520–600 nm	500–620 nm	500–620 nm
Red	630–690 nm	590–710 nm	590–710 nm
Near Infra Red	760–900 nm	740–940 nm	740–940 nm
Sun azimuthal angle	157.3°	147.3°	154.3°
Sun elevation angle	33.1°	66.2°	59.3°
View azimuthal angle	42.5°	179.9°	180.0°
View zenithal angle	20.0°	24.8°	13.5°
Study site	Nezer	Tagon	Tagon

related to each forest variable. Both single and multiple linear regressions are processed. For multiple linear regression, a new variable subset selection technique, we called Random PRiF, is proposed. The whole process will be detailed here after.

3.1. Image features

In our approach, two first order texture features (mean and variance, called var) and eight second order GLCM texture features were used (Haralick et al., 1973; Conners et al., 1984). The latter are considered as particularly relevant for forest applications (Franklin et al., 2001; Coburn and Roberts, 2004; Kayitakire et al., 2006; Murray et al., 2010; Wunderle et al., 2007; Ozdemir and Karnieli, 2011; Castillo et al., 2010; Chehata et al., 2014; Wood et al., 2012). They are described in Table 3.

All texture measures were calculated from panchromatic (Pan) and from four multispectral (MS) bands: Blue (B), Green (G), Red (R), Near Infra Red (NIR), with various combinations of GLCM parameter values. These parameters are: radius of the moving window r (window size = $2r + 1$), displacement d, orientation o and quantification level number *nbbin*. A limited parameter range was fixed so as to make a trade-off between computational costs and effectiveness. The ranges of r and d were defined according to the spatial resolution and are given in Table 4. The quantification level *nbbin* was set to 8 as commonly used in remote sensing literature.

3.2. Test frames

Our texture feature selection process was automatically performed on test frames gathering small image samples of all stands where the forest variables were measured in field. This method made it easy to explore Haralick's texture features across a wide range of parameter values (GLCM parameters). Indeed thousands

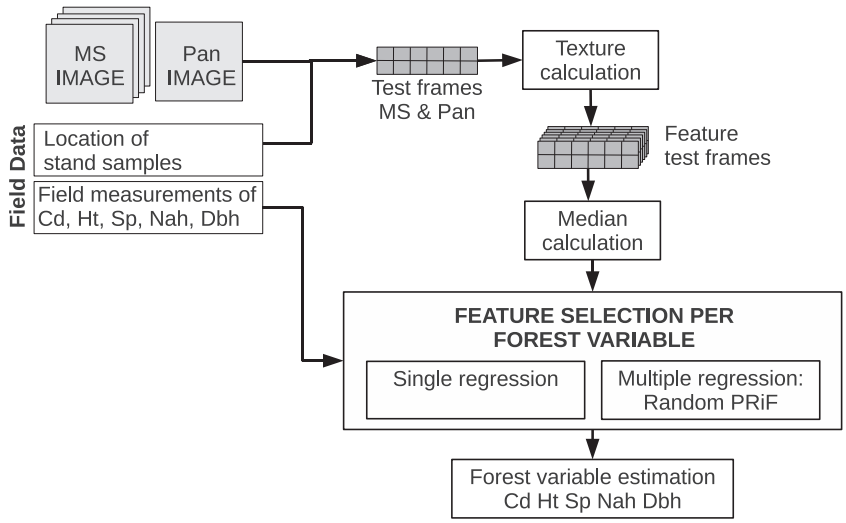


Fig. 3. A Global framework of forest variable retrieval.

Table 3
GLCM texture features description.

Feature	Description	Formula
Energy (energy)	High when the GLCM is locally homogeneous	$\sum_{i,j} g(i,j)^2$
Entropy (entropy)	Measure of the randomness of the intensity distribution	$-\sum_{i,j} g(i,j)\log_2 g(i,j)$
Correlation (corr)	Linear dependency on neighboring pixels values	$\sum_{i,j} \frac{(i-\mu)(j-\mu)g(i,j)}{\sigma^2}$
Inverse Difference (invdif)	A measure of homogeneous pixel values across an image, inversely related to Inertia	$\sum_{i,j} \frac{1}{1+(i-j)^2} g(i,j)$
Inertia (inertia)	High values are indicative of high grey level values off the matrix diagonal	$\sum_{i,j} (i-j)^2 g(i,j)$
Cluster Shade (clushade)	Higher order moment	$\sum_{i,j} ((i-\mu)+(j-\mu))^3 g(i,j)$
Cluster Prominence (cluspro)	Higher order moment	$\sum_{i,j} ((i-\mu)+(j-\mu))^4 g(i,j)$
Haralick's Correlation (harcorr)	Correlation measure	$\frac{\sum_{i,j}(i,j)g(i,j)-\mu_t^2}{\sigma_t^2}$

with:

- $g(i,j)$ a normalized GLCM matrix element.
- $\mu = \sum_{i,j} i \cdot g(i,j) = \sum_{i,j} j \cdot g(i,j)$ is the pixel values weighted mean.
- $\sigma = \sum_{i,j} (i-\mu)^2 \cdot g(i,j) = \sum_{i,j} (j-\mu)^2 \cdot g(i,j)$(due to matrix symmetry) is the pixel values weighted variance.

Table 4
Values of the spatial GLCM parameters.

Data	Radius r (pixel)	Displacement d (pixel)	Orientation o (degree)
Pan	5–25 step: 5	1–10 step :1	0–135° step: 45°
MS	3–12 step: 3	1–4 step:1	0–135° step: 45°

of descriptors were calculated given the large number of possible combinations of parameter values. The proposed feature selection scheme assesses the capability of diverse multiple features descriptors to retrieve the forest structure with minimal computational costs. Fig. 4 exhibits the few image samples used to build the test frames. Each image sample covers a square area centered at a field measurement plot (cf. Section 2.2) with a slightly larger size. Its width is 200 pixels in Pan and 50 pixels in MS (i.e. ~120 m) on Nezer site. It is 120 pixels in Pan and 30 pixels in MS for Tagon site (i.e. ~60 m).

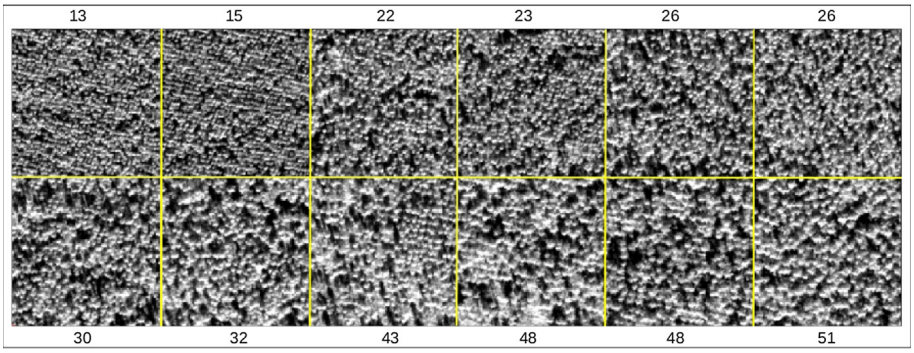
The maximal and minimal values of radiometry observed on the whole test frame were used to define the GLCM quantification levels in each spectral band of each image in order to guarantee a radiometry resolution for forest pixels in concordance with the used number of levels (nbbin = 8). For each sampled plot, the median value of each tested image feature was calculated and used as a sample centroid in the next step of selecting the most relevant features for retrieving the forest variables. The median value was preferred to the mean value, which is commonly used in remote sensing for this matter, in order to overcome smoothing effects.

3.3. Automated feature selection using linear regression

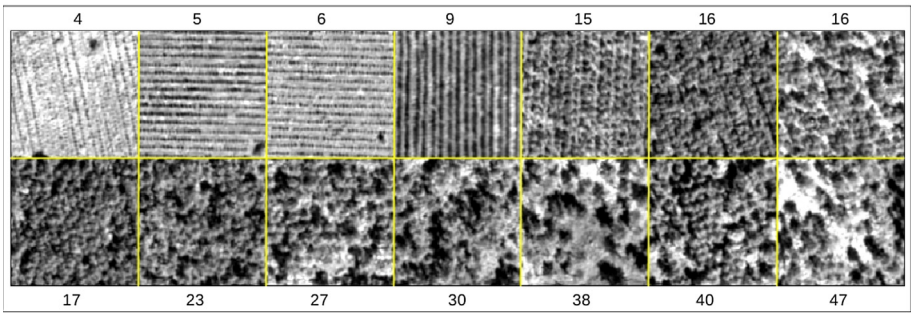
The selection of the most relevant texture features for retrieving each of the five forest variables is based on statistical modeling of forest variable as linear function of one or multiple image features. Modeling is based on linear regression. Single linear regression assumes that only one texture feature is sufficient to estimate the considered forest variable and thus exploits only one spectral band and only one spatial resolution. Multiple linear regression is applied on both Pan and MS texture features and thus allows to assess a multi-scale and multi-resolution modeling. Thus better accuracies are expected but multicollinearity effects have to be taken into account carefully.

3.3.1. Single variable linear regression

The linear regressions were performed between all forest structure variables and all texture features, derived from Pan and MS



(a) Nezer site (12 samples, image sample dimensions : 124*124m)



(b) Tagon site (14 samples, image sample dimensions : 60*60m)

Fig. 4. Subset of image samples in panchromatic band for both sites illustrating the variability in forest structure of the sampled stands. The samples are sorted by increasing tree age.

bands with different parameterizations. The determination coefficient R^2 was used as criterion to identify the best texture feature in Pan and in MS datasets for each date and each site. The processing was automated to select the texture feature leading to the highest R^2 with the forest variable. The accuracy of the linear modeling was estimated with the RMSE (root mean square error) calculated by the leave-one-out cross validation technique (named here LOO RMSE).

3.3.2. Multiple variable linear regression

Multiple linear regression has been applied in literature using SPOT-5 multispectral information only (initial bands, derived vegetation indexes and GLCM texture bands) (Castillo et al., 2010). Wunderle et al. (2007) used image texture derived from pan-sharpened SPOT imagery and stepwise regression analysis to model a stand structural index. Ozdemir and Karnieli (2011) modeled forest structure parameters as function of WorldView2 spectral texture measures using stepwise multi-linear regressions. Our approach explores the potential of combining Haralick's features derived from both Pan and MS bands, thus using jointly different spatial and spectral resolutions. In addition, the underlying automated processing allows to optimize GLCM's spatial parameters while exploring various texture feature combinations thus allowing multi-scale modeling. The multiple regressions have to be thus performed on thousands of descriptors, many of which being likely correlated. The main difficulty in multiple linear regression is minimizing the multicollinearity of the feature subset solutions in order to generate stable models and avoid over-fitting. This is generally done using the Variance Inflation Factor (VIF), a good multicollinearity detector which is defined as follows for any variable j used as regressor in multiple regression:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{2}$$

where R_j^2 is the determination coefficient obtained by regressing the jth regressor on the remaining regressors. A variance inflation factor exists for each of the j regressors in a multiple regression model. The higher VIF_j, the higher the collinearity between variable j and the other variables. In other words, VIF_j equals to 1 if there is no correlation between them. VIF values exceeding 4 warrant further investigation, while VIF values exceeding 10 are signs of serious multicollinearity requiring correction (Castillo et al., 2010; Haan, 2002; Kutner et al., 2004).

A common prediction criterion involved in multilinear regression is *PRESS* (Predicted Residual Error Square Sum) statistic (Allen, 1974) defined as:

$$PRESS = \sum_{i=1}^{n} (y_i - \hat{y}_i(-i))^2 \tag{3}$$

where $\hat{y}_i(-i)$ is the prediction for observation i using the model calculated from the $n - 1$ others.

PRESS is equivalent to a statistic of Leave-One-Out cross validation (LOO RMSE = $\sqrt{(PRESS/n)}$).

Various methods for feature subset selection exist using multiple regressions. In previous work (Beguet et al., 2012), the performances of three methods were compared: the classic step-by-step forward method, a LASSO (Least Absolute Shrinkage and Selection Operator) method, and a forward stepwise approach based on LARS (Least Angular Regression) (Efron et al., 2004). The latter is a newer forward stepwise method. Unlike the classic Forward method which incrementally adds the variable leading to the highest F-statistic, the LARS algorithm adds step-by-step the variable that better explains the current residuals. LARS stepwise turned out the most effective subset selection method, with both low *PRESS* and low *VIF* (below the critical value of 4 for each regressor). The main drawbacks of these step-by-step variable selection techniques are (1) they provide suboptimal solutions due to the underlying incremental optimization scheme and (2) the relative importance of

the selected variables are highly imbalanced and biased towards the first variable.

To alleviate the weakness of traditional step-by-step variable selection, we propose a new variable subset selection scheme, that we called "Random PRiF". Unlike LARS stepwise which is an incremental method (keeping the first best feature and adding some complementary ones) and hence suboptimal, our method is global and based on a simple yet statistically robust optimization scheme: random sampling (with replacement) in feature space. The main idea is to assess, through an iterative process, a large number of possible subsets of image features chosen randomly (with replacement) among the whole range of descriptors, by minimizing a new statistical criterion which appropriately combines both *PRESS* and *VIF* criteria. The aim is to automatically find a feature subset with a strong predictive performance while keeping a low *VIF*. The criterion we named *PRiF* is defined as follows:

$$PRiF = PRESS * \frac{\exp(V)}{V^2} \tag{4}$$

where V is the maximum of the m-component *VIF* vector associated to the m selected features.

The use of a function of *VIF* in *PRiF* that has a hyperbolic behavior favors feature subsets with a good prediction performance while penalizing those with high values of *VIF* in order to achieve both accuracy and stability in resulting models. Our technique provides optimal variable subset solutions whose relative importances are significantly more balanced than their step-by-step solutions counterpart.

Though inducing significantly higher computational costs than step-by-step approaches, the latter faces much higher risks of falling into local optima, leading to poor models. In Random PRiF, the higher the number of iterations of the random sampling-based procedure, the higher the achievable prediction accuracy and the lower the risk of a suboptimal solution. A comparative analysis is conducted in the following section between Random PRiF and two well-established methods: PCA (Principal Component Analysis) dimensionality reduction method and LARS stepwise. PCA has been chosen as a reference method without multicollinearity, i.e. *VIF* = 1 between the principal components used as regressors. LARS stepwise, a recent state-of-the-art variable subset selection method, is considered for performance comparison because it led to the best results in a previous comparative analysis (Beguet et al., 2012).

4. Results and discussion

The image features calculation was processed with the Orfeo-ToolBox library (OTB) http://www.orfeo-toolbox.org/otb/, the statistical analysis was carried out on R software. The whole processing chain was managed using Python programming language.

4.1. Single feature solutions

The single regressions were processed separately on Pan and MS datasets for both sites. Due to the diversity of the tested parameter values, 1730 and 2368 texture feature descriptors were tested for Pan and MS respectively. For each forest variable, numerous significant linear relationships with image features were found (considering a p-value <0.01 for R^2). Features with the highest R^2 are presented in Table 5 for Pan and MS datasets on Nezer and Tagon sites respectively. For each forest variable, the best PAN and MS features are presented with the corresponding parameterization and spectral band.

For Nezer site (Table 5), all considered forest variables were well retrieved by image texture features (R^2 varies from 0.89 to 0.98). The performances are mostly similar for Pan and MS on this reduced set of stands (n = 12). The texture features calculated on the Green band appeared to be the most effective MS ones. Second order texture features were clearly the most descriptive of the forest structure. In this case, inertia was the most frequently selected texture feature, for both resolutions and with various parameterizations. It is interesting to notice that for Pan features, window radius parameter varied from 5 to 25 pixels depending on the retrieved forest variable. This confirms that no a priori presumption could be made on GLCM parameters, thus justifying a wide-ranging and automatic search of their optimal values for the forest variable retrieval task.

For Tagon site (Table 5), the single regressions were processed for both dates: June and August. However due to paper length restriction, only June results are presented. The quality of predictions is lower than the one observed for Nezer. R^2 varies from 0.32 (Nah) to 0.73 (Sp) in June and from 0.34 (Nah) to 0.78 (Sp) in August and LOO RMSE is always higher. Similarly to Nezer site, the Green band provides the best single MS-based models. For both dates over Tagon, MS features appeared to be more relevant for forest structure estimation than PAN ones. Besides, the variance, a first order feature, gives some best results. Many factors such as acquisition conditions, sun position, phenology stage and spectral properties of the understory may influence this result. For instance, since Tagon site is characterized by a high variety of tree row orientations, the variance, which is insensitive to image texture anisotropy, could appear more robust to orientation variations than second order features. Similar values for R^2 and RMSE were achieved on June and August images, though slightly better for the August image, but with different best performing image features. Indeed, the image texture of the same stands varies probably between both dates due to small differences in phenology, sun and view angles. For instance, the spatial heterogeneity of the understory reflectance was probably slightly smaller in August due to the phenological stage (the leaf growth ended in August for all species unlike in June), the solar position induced slightly larger shadows on ground in August, also a better visibility of the tree dimensions

Table 5
The best performing single MS and Pan image features for retrieving each forest variable. Nezer site, October image. Tagon site, June image.

Forest var.	Band	Feature	r	d	o	R^2	p Value	LOO RMSE
NEZER								
Cd	B	cluspro	3	1	0	0.970	<10e−06	0.30
	Pan	inertia	10	5	45	0.979	<10e−06	0.25
Sp	G	inertia	3	2	0	0.961	<10e−06	0.37
	Pan	inertia	5	1	45	0.944	<10e−06	0.40
Ht	G	inertia	6	2	0	0.946	<10e−06	1.5
	Pan	inertia	25	1	45	0.905	<10e−06	1.6
Nah	B	corr	3	1	45	0.891	<10e−06	140
	Pan	inertia	5	1	135	0.909	<10e−06	37
Dbh	G	inertia	12	1	45	0.938	<10e−06	0.029
	Pan	inertia	10	5	45	0.927	<10e−06	0.029
TAGON								
Cd	G	var	6			0.650	<10e−06	1.397
	Pan	corr	10	2	45	0.505	<10e−06	1.659
Sp	G	var	3			0.737	<10e−06	0.979
	Pan	var	25			0.572	<10e−06	1.242
Ht	NIR	invdif	12	3	135	0.635	<10e−06	4.450
	Pan	corr	15	3	45	0.550	<10e−06	4.993
Nah	NIR	corr	12	1	135	0.445	<10e−06	682.112
	Pan	corr	10	2	45	0.321	<10e−06	716.776
Dbh	G	var	6			0.650	<10e−06	0.086
	Pan	corr	10	2	45	0.533	<10e−06	0.099

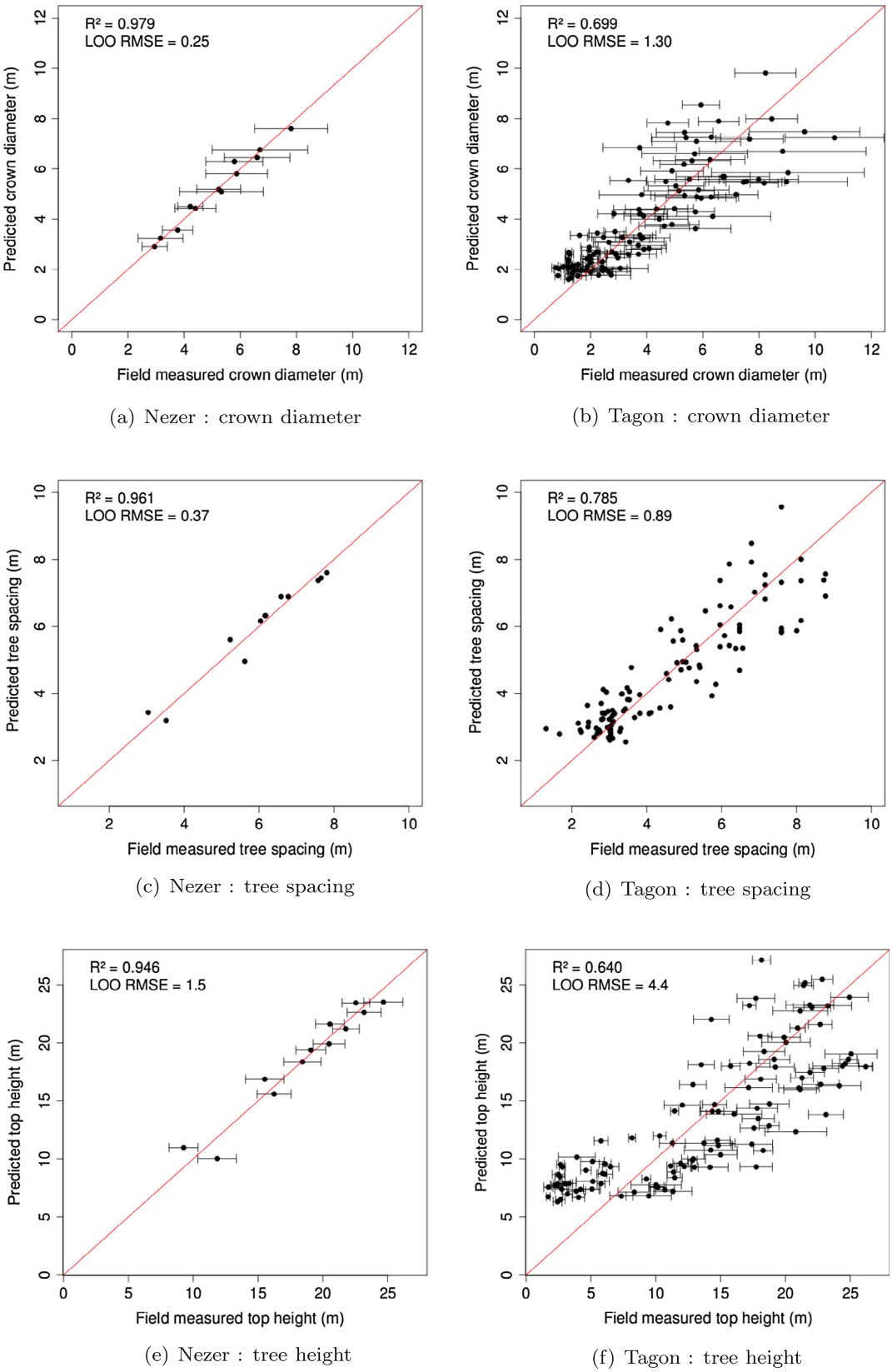


Fig. 5. Field measured forest variables vs. image derived estimates from best single variable models. Left: Nezer site, October image. Right: Tagon site, August image. Bars correspond to ±1 standard deviation.

was probably obtained with a lower view zenith angle in August. In spite of these small differences, the image texture was very informative on the forest structure for both dates. The performances of the best image feature subsets are relatively good and similar for both dates, thus seeming to be rather robust to image acquisition conditions.

For both sites and dates, the best prediction performances were obtained for tree spacing (*Sp*), crown diameter (*Cd*) and tree height (*Ht*) which are the forest variables that directly impact the image texture. However the trunk diameter *Dbh*, which has no direct effect on texture, also showed high R^2 values close to those obtained for *Cd*. This comes from the strong linear allometric relationships between both forest variables (see Fig. 2) already demonstrated for maritime pine (e.g. Porté et al. (2000)). The tree density *Nah* showed the lowest linear correlations with texture features unlike *Sp* which is calculated as a non-linear transform of *Nah* (*cf* Eq. (1)).

Fig. 5 compares the values measured on field and their estimates predicted from the best single variable regression model for *Cd*, *Sp* and *Ht*. No outliers are noticeable on residuals. The presumption of linear relationships for the three forest variables are clearly confirmed for the Nezer site, but the small number of observations does not allow a reliable interpretation of the results. The Tagon site is more informative since it provides a higher structure variability of maritime pine forest stands.

First, the variance (see standard deviations in Fig. 5) of field measured variables has to be considered when interpreting the quality of predictions. A heteroscedastic behavior is observed, more pronounced for crown diameter *Cd* than for tree height *Ht*. This highlights the fact that the older the stand, the higher the intra-stand variance of *Cd*. Indeed, with an increasing age and spacing between trees, the tree spatial distribution becomes more heterogeneous, thus inducing local variations in crown growth (due to tree competition) and thus in *Cd*. This also points out the uncertainty on *Cd* field measures. The higher the trees, the more difficult to measure the crown diameter.

Secondly, the relationships between measured and predicted values are not clearly linear except for *Sp* (Fig. 5(d)). The predicted values for *Cd* and *Ht* were overestimated for the youngest stands (Fig. 5(b) and (f)). For instance, for field measured *Ht*<5 m, the predicted values varies from 5 to 10 m. Moreover, for the three forest variables, the residual variance is proportional to field measured variable values. Thus using a single regression model is not really suitable for estimating a forest structure variables.

Finally these preliminary results showed the potential of an automated selection of the most informative texture descriptors (i.e. the best Haralick feature parameterization) to estimate forest structure variables.

4.2. Multiple feature solutions

For Nezer site, the number of texture features per regression was fixed to 3 using Mallow's Cp curve (Beguet et al., 2012). It was set to 5 for Tagon, as a trade-off between parsimony and complementarity of information by various texture features. The maximal iteration number used to obtain the best solution was set experimentally to 1 million for both sites.

4.2.1. Performance analysis of feature selection methods

The three competing feature selection methods (PCA dimensionality reduction, LARS stepwise and Random PRiF) were compared, on both sites, for all forest variables (5) and using three image datasets per date: Pan alone (texture calculated on Panchromatic band), MS alone (texture calculated on multispectral bands), and MS-Pan which is composed of all features from both resolutions (hereinafter called multi-scale MS-Pan dataset). Hence, 30 different experiments (2 sites $*$ 3 image dataset types $*$ 5 forest variables) comparing the three methods were repeated. For Tagon, only August image data was tested. Table 6 illustrate some of the results by showing the performances (R^2, PRESS, VIF, LOO RMSE) of the best performing feature subsets obtained from each of the three methods on the MS-Pan dataset of Nezer site and on the Pan dataset of Tagon site.

A voting procedure was applied on the 30 experiments to identify which feature selection method was the most relevant to predict forest variables using as criterion the smallest LOO RMSE. Our Random PRiF feature selection method got the best overall prediction performance in almost 77% of cases, followed by the LARS stepwise. The PCA reduction is the worst method in all cases (*cf.* Table 6). These results confirm the relevance and the effectiveness of the introduced PRiF criterion within a random global search context. In addition, the collinearity, which is a critical issue in multilinear regression, was low with maximum *VIF* values lower than the established critical threshold of 4 (Castillo et al., 2010) for all best solutions of both LARS stepwise and Random PRiF (*cf.* Table 6). As for PCA, which builds a new feature space composed of uncorrelated variables (of decreasing importance) and hence ensures by definition no collinearity, VIF is not a relevant criterion (VIF = 1 for all variables of a given subset). The results show that our Random PRiF method has a good behavior in terms of model stability and reliability, exhibiting maximal *VIF* values lower than 4 in all cases. LARS stepwise generally outperforms Random PRiF with respect to *VIF* criteria while it is the opposite with respect to prediction quality criteria (PRESS and LOO RMSE). But, this is the very essence of our novel variable selection criterion which was built so as to increase prediction accuracy while keeping collinearity reasonably low. We deliberately favors PRESS over *VIF* criterion because the former has a priori more impact on multi-linear overall model performances than the latter. PCA obviously provides the lowest possible *VIF* values but this never allowed this well-known method to outperform the two other assessed methods in all the 30 forest variable retrieval carried out experiments.

Besides, our Random PRiF method provided more balanced texture feature subsets than LARS stepwise with regard to the "variable importance" statistics. Variable importance is defined as the contribution of each regressor in the obtained model ; it was estimated through LMG statistic (Gromping, 2006). This is another appealing feature of our method that was also designed to alleviate some weaknesses of traditional step-by-step variable subset selection methods such as LARS stepwise (Efron et al., 2004) which provide suboptimal and highly imbalanced (large bias towards the first selected feature) solutions. Our variable subset selection methods relies on a simple and statistically robust optimization scheme: random sampling in feature space.

Table 7 outlines the importance of each texture feature in the selected subsets by both LARS stepwise and RandomPRiF methods for Nezer site. Random PRiF method led to significantly lower maximal importance values than LARS stepwise which exhibits a very large bias towards the first feature. For instance, for *Ht* variable, Random PRiF provides a balanced subset where the three texture features are roughly of same importance while LARS stepwise leads to a completely unbalanced feature subset. Indeed, the latter method is based on an incremental optimization scheme which is not the optimal way to combine features in a multiple regression framework. Though leading to suboptimal solutions, this step-by-step method has an advantage over our method : its lower computational complexity. This balanced behavior of Random PRiF method was clearly observed on Nezer site as well as on Tagon site though less dramatically. Thus the random global optimization approach used in Random PRiF helps to optimize the complementarity between features. It also leads to higher diversity in texture information capturing the essence of forest structure as opposed to step-by-step methods which more or less tend to reduce the potential of multiple-feature models to single-feature ones.

Finally, the results of the different experiments previously mentioned were summarized per forest variable and per site, a sample of which composed of the best performing pair (dataset/feature selection method) is presented in Table 8. It can be noticed that, the multi-scale MS-Pan dataset generally leads to the best models.

Table 6
Feature selection methods performance comparative analysis : statistics criteria of best feature subset using the MS-Pan dataset. October image on Nezer site (3 features subset), August image on Tagon site (5 features subset).

Forest var.	Selection method	multiple R^2	PRESS	Max(VIF)	LOO RMSE
NEZER					
Cd	PCA	0.917	5.458	1.000	0.67
	LARS stepwise	0.995	0.190	1.288	0.13
	Random PRiF	0.996	0.168	2.421	0.12
Sp	PCA	0.906	6.149	1.000	0.72
	LARS stepwise	0.994	0.535	1.025	0.21
	Random PRiF	0.996	0.233	2.734	0.14
Ht	PCA	0.903	66.034	1.000	2.3
	LARS stepwise	0.992	5.114	1.070	0.7
	Random PRiF	0.994	2.791	2.519	0.5
Nah	PCA	0.871	464568	1.000	197
	LARS stepwise	0.998	8154	1.287	26
	Random PRiF	0.997	5582	3.404	22
Dbh	PCA	0.892	0.029	1.000	0.049
	LARS stepwise	0.987	0.003	1.226	0.016
	Random PRiF	0.993	0.001	1.270	0.009
TAGON					
Cd	PCA	0.534	331.114	1.000	1.73
	LARS stepwise	0.752	180.304	1.357	1.27
	Random PRiF	0.800	129.015	2.293	1.08
Sp	PCA	0.548	214.820	1.000	1.39
	LARS stepwise	0.765	116.219	2.001	1.02
	Random PRiF	0.808	84.263	1.860	0.87
Ht	PCA	0.556	2904.690	1.000	5.1
	LARS stepwise	0.837	1144.369	2.052	3.2
	Random PRiF	0.851	1032.486	1.666	3.1
Nah	PCA	0.225	66281517	1.000	773
	LARS stepwise	0.523	51588535	1.408	682
	Random PRiF	0.500	46197877	1.748	645
Dbh	PCA	0.541	1231.149	1.000	0.105
	LARS stepwise	0.818	475.571	3.817	0.065
	Random PRiF	0.823	461.019	2.132	0.064

Table 7
Variable importance in the subsets of 3 features selected by LARS stepwise and Random PRiF methods, Nezer site, MS-Pan dataset.

Forest var.	Texture variable importance	
	LARS stepwise	Random PRiF
Cd	0.921	0.752
	0.011	0.272
	0.066	0.220
Sp	0.959	0.445
	0.035	0.442
	0.004	0.111
Ht	0.975	0.372
	0.018	0.326
	0.006	0.300

Table 8
Best multiple-variable solutions on Nezer and Tagon (August) sites with respect to both feature selection method and dataset type.

Forest var.	Site	Dataset	Feature selection method
Cd	Nezer	MS-Pan	Random PRiF
	Tagon	Pan	Random PRiF
Sp	Nezer	Pan	Random PRiF
	Tagon	MS-Pan	LARS Stepwise
Ht	Nezer	MS	Random PRiF
	Tagon	MS-Pan	LARS Stepwise
Nah	Nezer	MS-Pan	Random PRiF
	Tagon	MS-Pan	Random PRiF
Dbh	Nezer	MS-Pan	Random PRiF
	Tagon	MS-Pan	LARS Stepwise

This highlights the potential of combining both spatial and spectral resolutions. Regarding the performances of feature selection methods over 30 experiments, Random PRiF appears as the best alternative. As a result from these findings, the following sections focus only on the solutions provided by Random PRiF feature selection method using MS-Pan datasets.

4.2.2. Contribution of multi-scale texture features

The best performing texture feature subsets obtained from MS-Pan dataset using random PRiF selection method are presented in Table 9 for Nezer and Tagon (only August image) sites, respectively. For both sites, the selected subsets combined in most cases both Pan and MS features with various parameterizations in window size (r), displacement length (d) and orientation (o). It can be noticed that the multi-linear regression combines various second order texture features with different parameterizations and thus overcomes the anisotropy issue (i.e. tree spatial distribution by rows). For instance, the best feature subset to retrieve *Nah* on Nezer site combines two resolutions and three different orientations (45°, 90° and 135°).

Another example is the best feature subset to retrieve *Cd* on Tagon site which has a higher forest structure variability than Nezer site. It is composed of two GLCM moments in Pan with the same window size (15 pixels) and different displacements (1 and 5 pixels) providing multi-scale analysis. In addition, three lower

Table 9
Multiple solutions combining panchromatic and multispectral data by Random PRiF, Nezer site with October image and Tagon site with August image.

Forest var.	Band	Feature	*r*	*d*	*o*	multiple R^2	VIF	LOO RMSE
NEZER								
Cd	B	cluspro	3	1	0	0.996	2.421	0.12
	Pan	corr	25	9	135		2.382	
	B	corr	9	2	0		1.039	
Sp	Pan	invdif	5	2	135	0.996	2.735	0.14
	Pan	clushade	20	7	135		1.998	
	G	cluspro	6	4	0		2.412	
Ht	R	corr	9	3	0	0.994	1.164	0.5
	R	cluspro	6	2	0		2.346	
	R	inertia	3	1	90		2.519	
Nah	NIR	corr	12	2	45	0.997	1.069	22
	R	energy	3	2	90		3.404	
	Pan	inertia	15	2	135		3.389	
Dbh	Pan	invdif	10	4	45	0.993	1.270	0.009
	R	corr	3	2	0		1.062	
	G	entropy	6	3	135		1.282	
TAGON								
Cd	Pan	harcorr	15	1	135	0.807	2.261	1.10
	Pan	cluspro	15	5	0		2.240	
	NIR	harcorr	12	5	0		2.140	
	NIR	clushade	3	2	45		1.009	
	NIR	inertia	12	3	135		2.128	
Sp	Pan	harcorr	20	8	0	0.819	1.127	0.84
	B	harcorr	9	2	0		1.918	
	R	cluspro	12	1	45		2.23	
	R	var	6				2.303	
	G	harcorr	3	1	45		2.100	
Ht	R	cluspro	9	1	0	0.831	1.607	3.1
	Pan	entropy	25	4	0		1.778	
	Pan	inertia	25	1	0		1.256	
	Pan	clushade	15	10	135		1.148	
	Pan	inertia	25	2	90		1.290	
Nah	NIR	harcorr	9	2	0	0.625	1.562	586
	G	clushade	3	1	0		1.107	
	G	inertia	6	5	0		2.106	
	NIR	corr	3	1	135		1.225	
	NIR	invdif	12	5	0		2.526	
Dbh	Pan	inertia	20	2	0	0.808	1.055	0.066
	G	var	6				1.192	
	G	corr	12	1	135		1.188	
	B	harcorr	12	1	90		2.021	
	G	harcorr	9	5	0		2.070	

resolution texture features, calculated on the NIR band, with different parameterizations, completed the subset.

While single feature solutions (*cf.* Section 4.1) were essentially first order variance features for Tagon site, which are insensitive to row direction variability, multiple feature solutions took advantage of the GLCM texture features combining them with different parameterizations automatically. Moreover, as previously observed for single solutions, the obtained multiple solutions, for both dates on Tagon site, are different but their prediction performances (R^2, LOO RMSE) are similar. Thus the proposed automated selection of best texture features allows to exploit the texture information related to the considered forest variable regardless of tree row direction variability and image acquisition conditions. Fig. 6 gives a comparison between the values measured in field and their estimates predicted from the best multiple variable regression models for *Cd*, *Sp* and *Ht*. The multiple variable models clearly reduce the residuals in comparison with single variable models (*cf.* Fig. 5) for both sites. The residuals are also distributed quite randomly around the line 1:1; reflecting a nearly constant residual variance whatever the value of the forest variable unlike single solutions. As a result, the individual prediction error was minimized with no noticeable over or underestimation over the variation range of each forest variable in comparison with the uncertainties on their values measured in field (see the length of bar indicating standard deviations).

In Table 10, the overall performances of single and multiple solutions in retrieving forest variables are compared according to LOO RMSE for both sites based on all available images (one date for Nezer and two dates for Tagon).

All prediction scores are improved with multiple solutions. On Nezer site, LOO RMSE was strongly improved for *Cd* by 52%, for *Ht* by 66% and for *Sp* by 62%. On Tagon site (August), which has a larger range of variation in forest variables and a larger forest structure variability between and within sampled stands compared to Nezer site, LOO RMSE was improved by 15.3% for *Cd*, by 29.5% for *Ht* and by 6.7% for *Sp*.

The gain in accuracy is better for *Dbh* (69%, 28%, 20% for Nezer, Tagon in June and in August, respectively) than for *Cd* (52%, 20%, 15%). Even if *Cd* and *Dbh* are highly correlated (*cf.* Fig. 2)) and *Cd* impacts directly the image texture unlike *Dbh*, this could be explained by a higher relative uncertainty in field measurement of *Cd* (see standard deviations in Fig. 6(a) and (b)). The gain is high for *Ht* (67%, 20%, 30% for Nezer, Tagon in June and in August, respectively), due probably to the greater complexity of the contribution of this variable to texture which is related to the length and direction of the crowns and the corresponding shadows as viewed

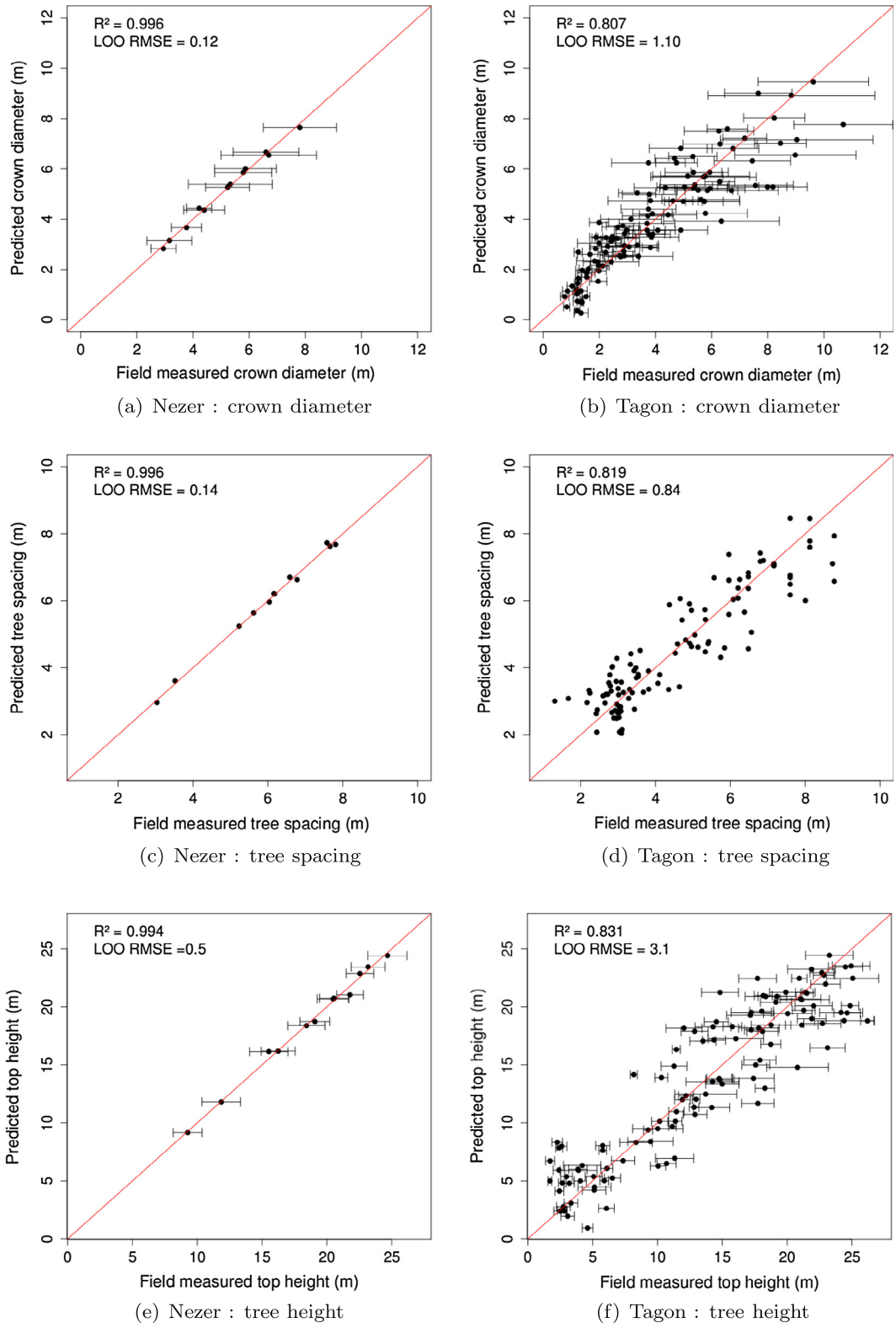


Fig. 6. Field measured forest variables vs. image derived estimates from best multiple variable models provided by Random PRiF. Left: Nezer site, October image. Right: Tagon site, August image. Bars correspond to ±1 standard deviation.

by the sensor. The gain is the lowest for *Sp* in Tagon (6% in June and 7% in August). The single solution accuracy was already high for Nezer with a limited number of sampled plots ($n = 12$) and the forest structure is more homogeneous within the sampled stands with a lower range of variation between them (age > 12 years). These conditions were optimal to produce a quasi-linear mono-variable model, which is likely to be unrobust for prediction over other stands.

Table 10
Performance comparison of single and multiple solutions obtained on the two sites from the MS-PAN dataset.

Forest var.	Site	LOO RMSE		Variable range
		Single	Multiple	
Cd (m)	Nezer	0.25	0.12	2.95–7.81
	Tagon (June)	1.40	1.12	0.77–10.69
	Tagon (August)	1.30	1.10	0.77–10.69
Sp (m)	Nezer	0.37	0.14	3.04–7.81
	Tagon (June)	0.98	0.92	1.31–8.77
	Tagon (August)	0.89	0.83	1.31–8.77
Ht (m)	Nezer	1.5	0.5	9.3–24.7
	Tagon (June)	4.5	3.6	1.7–26.2
	Tagon (August)	4.4	3.1	1.7–26.2
Nah (tree/ha)	Nezer	37	22	189–1253
	Tagon (June)	682	640	150–6729
	Tagon (August)	677	586	150–6729
Dbh (m)	Nezer	0.029	0.009	0.15–0.46
	Tagon (June)	0.086	0.062	0.02–0.56
	Tagon (August)	0.081	0.065	0.02–0.56

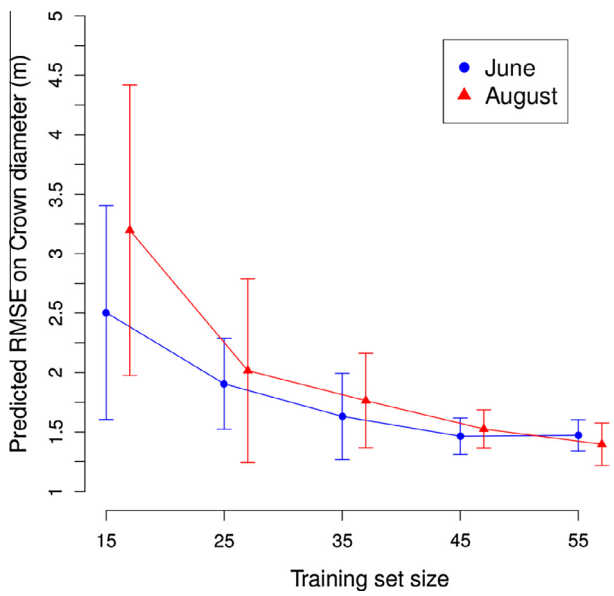


Fig. 7. Average RMSE distribution as function of the number of samples in training set. Results for Cd from same test set ($n = 40$) and with 30 iterations per different size of the training set. Bar = ±1 standard deviation.

Finally, the best LOO RMSE observed for Tagon using multiple solutions are very satisfactory for *Cd* and *Sp*, since they are very close to the expected accuracy with respect to image resolution: LOO RMSE ~1.1 m for *Cd* and ~0.8 m for *Sp*, i.e. ~ twice the resolution in Pan (0.5 m). Thus, for the youngest stands with tree size or spacing smaller than 1 m, it would be hard to find a model providing a good accuracy.

These results show that the multiple solutions are suitable to produce forest variable estimates. The texture feature subset selected to estimate each forest variable and the accuracy of the derived predictions were inherent to the studied site and the used image (sensor, date). A model that was determined from a given field-measured dataset and a given image gave poor estimation results when applied to another site or another image. This can be explained by various uncontrolled parameters such as seasonal variations (i.e. phenology stages, solar elevation range) or the viewing angle during image acquisition, which affects the interactions of radiation with a given forest canopy as measured by the sensor, and thus image texture, as previously mentioned.

The differences in performance between the images acquired on Tagon in June and August were quite small. However, this study did not consider images acquired on the same site with very contrasting various conditions (winter to summer, nadir to very oblique viewing). Further work should be carried out to identify the best image conditions to ensure a satisfactory prediction performance, i.e. a texture that is strongly related to the tree canopy structure.

4.2.3. Assessment of prediction performance

This last section focuses on the validation of our method for modeling a given forest variable using a multi-variable solution. This assessment was made on Tagon site where various field observations are available ($n = 111$) and are assumed to be representative of the forest structure variability in maritime pine stands over the study area.

The field samples were split into two independent datasets: a training set to find the best regression model using Random PRiF method on the MS-Pan image data and a test set to calculate the accuracy (i.e. RMSE) of the forest variable predicted from this model. The prediction robustness was assessed by repeating the modeling step on different training sets of increasing size. Thus the initial field samples were split into two subsets of same distribution (the samples were sorted with respect to their crown diameter and were assigned alternatively either as training or as test sample). The first subset contained 71 samples from which all training sets were randomly sampled with replacement. The same sampling was drawn for both dates. The second subset which contained the remaining 40 samples was the only one test set. The RMSE was computed from the predictions calculated on test set while increasing the size of the training set from 15 to 55 samples. The RMSE was averaged over 30 runs.

Fig. 7 illustrates the results obtained on June and August images for *Cd* (crown diameter). It shows that the average RMSE decreases significantly while increasing the training set size. Average RMSE decreases by approximately 40% from 15 observations within the training set to 45 observations (i.e. from 2.5 m to 1.5 m and from 3.2 m to 1.4 m for June and August images, respectively). Then it reaches a plateau with a value close to the LOO RMSE previously calculated (~1.1 m for both dates). On the plateau, the maximal standard deviation is approximately 0.50 m. These results show the minimum number of samples required to take into account the variability of the forest structure that is common in the Landes de Gascogne forest and to ensure a good predicting performance of main forest variables. For Tagon site, using an equal number of 40 samples in both training and test sets is a good trade-off to guarantee for both dates a model estimating *Cd* with a maximum error of ~2 m. With such a model, it would be possible to identify several classes of crown diameters ranging from 1 m or less to about 11 m. Nevertheless, further work should be undertaken with more in-field samples in order to refine the estimates of the expected maximal errors on the different forest variables.

5. Conclusion

This work provided an automated method to retrieve forest structure variables using texture information extracted from VHR optical satellite images. An automated feature selection process based on the use of image test frames of forest samples was investigated. This method allows a fast assessment of thousands of descriptors, exploring a wide range of parameter values, to optimize spatial texture parameter settings. The best image texture features are selected inherently leading to multi-scale texture

descriptors via statistical modeling. Five typical forest structure variables were successfully modeled. Our forest variable retrieval undergoes both multi-resolution and multi-scale texture analysis, combining panchromatic and/or multispectral features with different spatial parameter settings. The multicollinearity problem, encountered in multi-linear regression, is addressed carefully using the VIF criterion. We propose a new variable subset selection scheme, we called Random PRiF, based on a global optimization via random sampling in feature space. This approach involves a new statistical criterion we named PRiF, combining both VIF and PRESS criteria. Comparing our method to two well-established feature subset selection methods (LARS stepwise and PCA), it turned out the most convincing, significantly improving the quality of estimation of all forest structure variables for both sites. As future work, other feature subset selection methods such as tree-based regression for instance should be tested.

Exploring texture feature combinations from both panchromatic and multispectral datasets leads to the best prediction performances. This approach could be extended to other applications involving VHR image texture. The whole protocol we have introduced can be easily applied to any other forest using site-specific test frames. Finally, the use of Pléiades images showed the great potential of this sensor for forest applications.

Future work should be conducted on various acquisition conditions on the same site (winter to summer, nadir to very oblique viewing) to identify the best image conditions to ensure a satisfactory prediction performance.

Acknowledgements

This research was funded with grants from Conseil Régional d'Aquitaine and CNES (Centre National d'Etudes Spatiales). The VHR images were provided by the Orfeo program set up by CNES in order to prepare the use of Pléiades images. The authors are very grateful to Bernard Issenhut (INRA, Unité expérimentale Forêt Pierroton) for the measurements in the field and the good quality of the provided forest data. Further thanks to Christian Germain from IMS Lab (Bordeaux), Jordi Inglada from CESBIO (Toulouse) and Jean-Pierre Wigneron from INRA (Bordeaux) for discussions.

References

Allen, D.M., 1974. The relationship between variable selection and prediction. Technometrics 16 (125), 125–127.
Barbier, N., Couteron, P., Proisy, C., Malhi, Y., Gastellu-Etchegorry, J.P., 2010. The variation of apparent crown size and canopy heterogeneity across lowland Amazonian forests. Global Ecol. Biogeogr. 19, 72–84.
Beguet, B., Chehata, N., Boukir, S., Guyon, D., 2012. Retrieving forest structure variables from Very High Resolution satellite images using an automatic method. ISPRS Ann. Photogramm., Remote Sens. Spatial Inform. Sci. I-7, 1–6.
Beguet, B., Boukir, S., Guyon, D., Chehata, N. 2013. Modelling-based feature selection for classification of forest structure using very high resolution multispectral imagery, SMC'2013, IEEE Int. Conf. on Systems, Man, and Cybernetics, Manchester, UK, pp. 4294–4299.
Boukir, S., Orny, C., Chehata, N., Guyon, D., Wigneron, J., 2013. Détection de changements structurels sur des images satellite haute résolution. Application en milieu forestier., Traitement du signal 6, 401–429.
Castillo, M.A., Ricker, M., Jong, B.H.J.D., 2010. Estimation of tropical forest structure from Spot-5 satellite images. Int. J. Remote Sens. 31 (10), 2767–2782.
Chehata, N., Orny, C., Boukir, S., Guyon, D., 2011. Object-based forest change detection using high resolution satellite images, In: PIA'2011, ISPRS Photogrammetric Image Analysis Conference, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich, Germany, vol. 38(3), pp. 49–54.
Chehata, N., Orny, C., Boukir, S., Guyon, D., 2014. Object-based change detection in wind-storm damaged forest using high resolution multispectral images. Int. J. Remote Sens. 35, 4758–4777.
Coburn, C.A., Roberts, A.C.B., 2004. A multiscale texture analysis procedure for improved forest stand classification. Int. J. Remote Sens. 25, 4287–4308.
Conners, R., Trivedi, M., Harlow, C., 1984. Segmentation of a high-resolution urban scene using texture operators. Comput. Vision Graphics Image Process. 25 (3), 273–310, ISSN 0734189X, doi:10.1016/0734-189x(84)90197-x.
Couteron, P., Pelissier, R., Nicolini, E.A., Paget, D., 2005. Predicting tropical forest stand structure parameters from Fourier transform of very high-resolution remotely sensed canopy images. J. Appl. Ecol. 42, 1121–1128.
Culbert, P.D. et al., 2009. The impact of phenoligical variation on texture measures of remotely sensed imagery. IEEE J. Sel. Topics Appl. Earth Observations Remote Sens 2 (4), 299–309.
Dell'Acqua, F., Gamba, P., 2003. Texture-based characterization of urban environments on satellite SAR images. IEEE Trans. Geosci. Remote Sens. 41 (1), 153–159.
Efron, B., Hastie, T., Tibshirani, I.J.R., 2004. Least angle regression. Ann. Statist. 32 (2), 407–499.
Feng, Y., Li, Z., Tokola, T., 2010. Estimation of stand mean crown diameter from high-spatial-resolution imagery based on a geostatistical method. Int. J. Remote Sens. 31 (2), 363–378.
Franklin, S.E., Maudie, A.J., Lavigne, M.B., 2001. Using spatial co-occurence texture to increase forest structure and species composition classification accuracy. Photogramm. Eng. Remote Sens. 67 (7), 849–855.
Gaetano, R., Scarpa, G., Poggi, G., 2009. Hierarchical texture-based segmentation of multiresolution remote sensing images. IEEE Trans. Geosci. Remote Sens. 47, 2129–2141.
Gomez, C., Wulder, M.A., Montes, F., Delgado, J.A., 2012. Modeling forest structural parameters in the mediterranean pines of central Spain using Quickbird-2 imagery and classification and regression tree analysis (CART). Remote Sens. 4, 135–139.
Gonzales, R.C., Woods, R.E., 2008. Digital Image Processing. Pearson Education Inc.
Gromping, U., 2006. Relative importance for linear regression in R: the package relaimpo. J. Statist. Softaware 17, 1–27.
Guyon, D., Riom, J., 1996. Estimation de caractéristiques forestières à partir d'images à haute résolution spatiale (Spot5). BUL.S.F.P.T. 141, 46–50.
Haan, C.T., 2002. Statistical methods in Hydrology, second ed. Iowa State University Press, Ames, Iowa.
Haralick, R.M., Shanmugan, K., Dinstein, I., 1973. Texture features for image classification. IEEE Trans. Syst. Man Cybern., 610–621.
Hyyppa, J., Hyypa, H., Inkinen, M., Engdahl, M., Linko, S., Zhu, Y.H., 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stands attributes. Forest Ecol. Manage. 128 (1-2), 109–120.
Kayitakire, F., Hamel, C., Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. Remote Sens. Environ. 102, 390–401.
Kutner, M.H., Nachtsheim, C.J., Neter, J., 2004. Applied Linear Regression Models, forth ed. McGraw-Hill Irwin.
Maillard, P., 2003. Comparing texture analysis methods through classification. Photogramm. Eng. Remote Sens. 69 (4), 357–367.
Manly, B.F.J., 1994. Multivariate Statistical Methods : A Primer. Chapman & Hall, London, UK.
Murray, H., Lucieer, A., Williams, R., 2010. Texture-based classification of sub-Antarctic vegetation communities on Heard Island. Int. J. Appl. Earth Observation Geoinformation 12, 138–149.
Ozdemir, I., Karnieli, A., 2011. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. Int. J. Appl. Earth Observation Geoinformation 13, 701–710.
Pacifici, F., Chini, M., Emery, W.J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. Remote Sens. Environ. 113 (6), 1276–1292.
Porté, A., Bosc, A., Champion, I., Loustau, D., 2000. Estimating the foliage area of Maritime pine (Pinus pinaster Ait.) branches and crowns with application to modelling the foliage area distribution crown. Ann. For. Sci. 57, 73–86.
Proisy, C., Couteron, P., Fromard, F., 2007. Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images. Remote Sens. Environ. 109, 379–392.
Regniers, O., DaCosta, J.-P., Grenier, G., Germain, C., Bombrun, L., 2013. Texture based image retrieval and classification of very high resolution maritime pine forest images. In: Proceedings IEEE International Geoscience and Remote Sensing Symposium (IGARSS).
Ruiz, L.A., Fdez-Sarria, A., Recio, J.A. 2004. Texture feature extraction for classification of remote-sensing data using wavelet decomposition: a comparative study. In: 20th ISPRS Congress.
Soe, W.M., Tyler, S.N.L.J.M., 2004. Wavelets for urban spatial feature discrimination: comparisons with fractal, spatial autocorrelation, and spatial co-occurrence approaches. Photogramm. Eng. Remote Sens. 70 (7), 803–812.
Song, C., Dickinson, M.B., Su, L., Zhang, S., Yaussey, D., 2010. Estimating average tree crown size using spatial information from Ikonos and QuickBird images: across-sensor and across-site comparisons. Remote Sens. Environ. 114, 1099–1107.
St-Onge, B.A., Cavayas, F., 1997. Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates. Remote Sens. Environ. 61, 82–95.
Trias-Sanz, R., Stamon, G., Louchet, J., 2008. Using colour, texture, and hierarchical segmentation for high-resolution remote-sensing. ISPRS J. Photogramm. Remote Sens. 63, 156–168.
Tuominen, S., Pekkarinen, A., 2005. Performance of different spectral and textural aerial photograph features in multi-source forest inventory. Remote Sens. Environ. 94, 256–268.

Tuttle, E.M., Jensen, R.R., Formica, V.A., Gonser, R.A., 2006. Using remote sensing image texture to study habitat use patterns: a case study using the polymorphic white-throated sparrow (Zonotrichia albicollis). Global Ecol. Biogeogr. 15, 557–559.

Van Coillie, F.M., Verbeke, L.P., Wulf, R.R.D., 2007. Feature selection by genetic algorithms in object-based classification of {IKONOS} imagery for forest mapping in Flanders, Belgium. Remote Sens. Environ. 110 (4), 476–487, ISSN 0034-4257.

Wolter, T., Townsend, P.A., Sturtevant, B.R., 2009. Estimation of forest structural parameters using 5 and 10 meter Spot-5 satellite data. Remote Sens. Environ. 113, 2019–2036.

Wood, E.M., Pidgeon, A.M., Radeloff, V.C., Keuler, N.S., 2012. Image texture as a remotely sensed measure of vegetation structure. Remote Sens. Environ. 121, 516–526.

Wulder, M.A., Drew, E.F.L., Franklin, S.E., Lavigne, M.B., 1998. Aerial image texture information in the estimation of Northern deciduous and mixed wood forest leaf area index (LAI). Remote Sens. Environ. 64, 64–76.

Wunderle, A.L., Franklin, S.E., Guo, X.G., 2007. Regenerating boreal forest structure estimation using Spot-5 pansharpened imagery. Int. J. Remote Sens. 28 (19), 4351–4364.

5.1.2 Sélection d'attributs pour la fusion de données optique et LiDAR

Approche en 2D

AIRBORNE LIDAR FEATURE SELECTION FOR URBAN CLASSIFICATION USING RANDOM FORESTS

Nesrine Chehata[1,2], Li Guo[1], Clément Mallet[2]

[1] Institut EGID, University of Bordeaux, GHYMAC Lab:
1 allée F.Daguin, 33607 Pessac, France.
[1] Institut Géographique National, MATIS Lab:
2-4 avenue pasteur, 94165 St Mandé, France.

Commission III - WG III/2

KEY WORDS: Lidar, Full-waveform, Classification, Feature selection, Random Forests, Urban scenes

ABSTRACT:

Various multi-echo and Full-waveform (FW) lidar features can be processed. In this paper, multiple classifers are applied to lidar feature selection for urban scene classification. Random forests are used since they provide an accurate classification and run efficiently on large datasets. Moreover, they return measures of variable importance for each class. The feature selection is obtained by backward elimination of features depending on their importance. This is crucial to analyze the relevance of each lidar feature for the classification of urban scenes. The Random Forests classification using selected variables provide an overall accuracy of 94.35%.

1 INTRODUCTION

Airborne lidar systems have become an alternative source for the acquisition of altimeter data providing unstructured 3D point clouds that describe the Earth's topography. The altimeter accuracy of a topographic lidar measurement is high (<0.1 m). Depending on the geometry of illuminated surfaces, several backscattered echoes can be recorded for a single pulse emission. Many authors have shown the potential of multi-echo lidar data for urban area analysis and building extraction (Sithole and Vosselman, 2004). 3D point cloud classification can be based on geometric and textural attributes (Matikainen et al., 2007). Other works include the lidar intensity (Charaniya et al., 2004) or combine lidar and multispectral data (Secord and Zakhor, 2007, Rottensteiner et al., 2005). Since 2004, full-waveform (FW) lidar systems have emerged with the ability to record the complete waveform of the backscattered 1D-signal laser pulse. In (Gross et al., 2007, Wagner et al., 2008), FW lidar features were used to detect vegetated areas. In urban scenes, the potential of such data has been barely investigated. For the analysis of laser scanner data, various classification techniques have been applied such as unsupervised classification by ISODATA (Haala and Brenner, 1999) and supervised classification as Bayesian networks (Stassopoulou et al., 2000), Dempster shafer fusion theory (Rottensteiner et al., 2005), Support Vector Machines (Secord and Zakhor, 2007, Charaniya et al., 2004, Mallet et al., 2008) or classification trees (Ducic et al., 2006, Matikainen et al., 2007).

In this work, we study different lidar features, multiecho and full-waveform to classify urban scenes into four classes: Buildings, vegetation, natural ground and artificial ground. Artificial ground gathers all kinds of streets and street items such as cars, traffic lights whereas the natural ground includes grass, sand, and bare-earth regions. No filtering is applied before the classification. The objective is to select the most relevant features for classifying urban scenes and to provide an accurate classification with a small number of features. We propose to achieve both objectives using Random Forests. It is an ensemble classifier based on decision trees. It returns good classfication results and provides also feature selection.
The paper is organized as follows. The lidar features will be detailed in Section 2. In Section 3, multiple classifiers are presented and especially Random Forests. The feature selection process is detailed in Section 3.3. Experimental results are then discussed in Section 4 and finally, conclusions are drawn in Section 5.

2 AIRBORNE LIDAR FEATURES

Multi-echo and full-waveform lidar data are available. The feature vector is composed of 21 components: 17 multi-echo and 4 full-waveform lidar features. The multi-echo lidar features are separated into height-based, echo-based, eigenvalue-based and local 3D-plane based features. The resulting feature vector fv for each site is given by:

$$\begin{aligned} fv = [&\Delta z\ \Delta z_{fl}\ \sigma_z^2\ \mathcal{C};\ N\ N_e\ ; \\ &\lambda_1\ \lambda_2\ \lambda_3\ A_\lambda\ P_\lambda\ S_\lambda\ L_\lambda; \\ &N_z\ \sigma_{N_z}^2\ \mathcal{R}_z\ D_\Pi;\ A\ w\ \sigma\ \alpha]^T \end{aligned}$$

All these features are computed using a volumetric approach within a local neighborhood $\mathcal{V}_P$ at each lidar point P. The local neighborhood includes all the lidar points within a cylinder, with a fixed radius, centered at the point. Lidar points are then projected into a 2D image geometry (0.5 m resolution). Feature images are obtained by computing, for each pixel, the mean corresponding value of the lidar points included in a 3×3 raster kernel (*cf.* Figure. 1). Table 1 summarizes input lidar features which are separated into five groups. Lidar features are detailed hereby.

2.1 Height-based lidar features

The first group is based on 3D point heights:

- Δz: Height difference between the lidar point and the lowest point found in a large cylindrical volume whose radius has been experimentally set to 15 m. This feature will help discriminating ground and off-ground objects (*cf.* Figure. 1(b)).
- Δz_{fl}: Height difference between First and Last pulses of the waveform of the current lidar point. It helps discriminating building roofs and ground.
- σ_z^2: The height variance. This feature has high values for vegetation (*cf.* Figure. 1(c)).

- $\mathcal{C}$: The local curvature is a discrete version of the Laplace operator. It is the maximum value of the gradient differences on heights, which are computed in four main directions based on a raster grid (Steinle and Vögtle, 2001). A 3×3 raster kernel has been chosen.

2.2 Echo-based features

- N: Total number of echoes within the waveform of the current lidar point. This feature is high for vegetation and building facades.
- N_e: Normalized number of echoes obtained by dividing the echo number by the total number of echoes within the waveform of the current lidar point. This feature highlights the vegetation since multiple reflections can occur on it (*cf.* Figure. 1(d)).

2.3 Eigenvalue-based lidar features

The variance-covariance matrix is computed within the local neighborhood $(V)_P$. The Eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$ are used as features. Besides, they provide additional features and help discriminating planes, edges, corners, lines, and volumes (Gross and Thoennessen, 2006). These features describe the spatial local distribution of the 3D points.

$$\text{Anisotropy} = A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1} \quad (1)$$

$$\text{Planarity} = P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1} \quad (2)$$

$$\text{Sphericity} = S_\lambda = \frac{\lambda_3}{\lambda_1} \quad (3)$$

$$\text{Linearity} = L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1} \quad (4)$$

λ_3 has low values for planar objects and higher values for non-planar objects (*cf.* Figure. 1(e)). The planarity feature shows high values especially for planar objects. Conversely, the sphericity feature gives high values for isotropic distributed 3D neighborhood.

2.4 Local plane-based lidar features

The planarity of the local neighboorhood will help discriminating buildings from vegetation. The local plane Π_P within $\mathcal{V}_P$ is estimated using a robust M-estimator with norm $L_{1.2}$ (Xu and Zhang, 1996).

- N_z: Deviation angle of Π_P normal vector from the vertical direction. This feature highlights the ground (*cf.* Figure.1(f)).
- σ_z^2: Variance of deviation angles within $\mathcal{V}_P$. It discriminates planar surfaces such as roads and building roofs from vegetation.
- $\mathcal{R}_\pi$: Residuals of the local plane estimated in a small vertical cylinder (0.5 m radius). Residuals $\mathcal{R}_\pi$ are calculated w.r.t the estimated plane as follows:

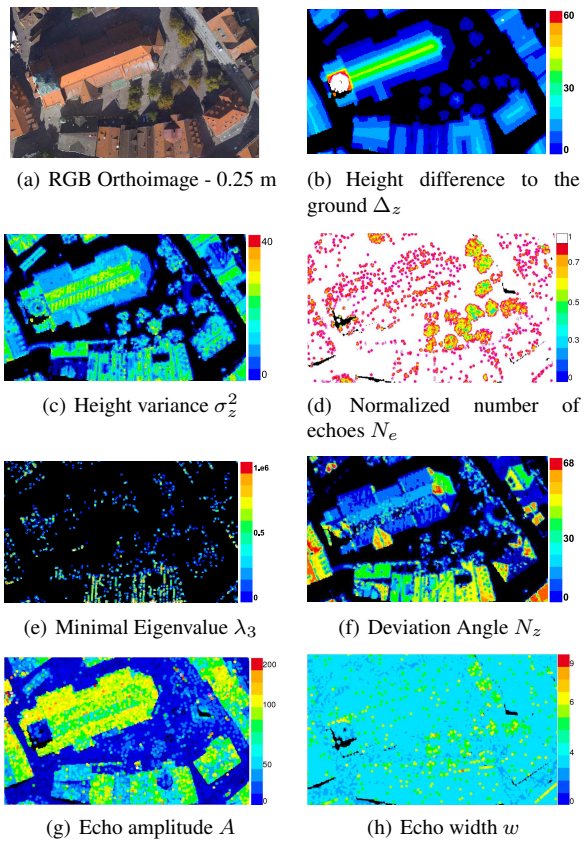
$$\mathcal{R}_\pi = \sum_{i \in \nu_P} \frac{(d_i)^l}{l} \quad (5)$$

  where d_i is the distance between the lidar point $i \in \nu_P$ and the plane. Here l=1.2. Residuals should be high for vegetation.
- D_π: Distance from the current point P to the local estimated plane Π_P.

2.5 Full-waveform lidar features

The remaining features are more specific to FW lidar data and are obtained by modelling the lidar waveforms. The amplitude and echo width are described in (Wagner et al., 2006).

- A: echo amplitude. High amplitude values can be found on building roofs, on gravel, and on cars. Asphalt and tar streets have low values. The lowest values correspond to vegetation due to a higher target heterogeneity and attenuation (*cf.* Figure. 1(g)).
- w: echo width. Higher values correspond to vegetation since it spreads the lidar pulse. A low width is likely to correspond to ground and buildings. However, it may increase with roof slope (*cf.* Figure. 1(h)).
- α: echo shape describing how locally spread the peak is. It is obtained from a waveform decomposition based on a generalized Gaussian modeling function (Chauve et al., 2007). The authors also show that very low and high shape values correspond mainly to building roofs and vegetation.
- σ: echo cross-section corresponds more or less to the peak energy. $\sigma = A \times w$. It is the basic quantity to describe the scattering of a wave by an object. The cross-section values are high for buildings, medium for vegetation and low for artificial ground.



(a) RGB Orthoimage - 0.25 m (b) Height difference to the ground Δ_z

(c) Height variance σ_z^2 (d) Normalized number of echoes N_e

(e) Minimal Eigenvalue λ_3 (f) Deviation Angle N_z

(g) Echo amplitude A (h) Echo width w

Figure 1: Orthoimage and some representative Lidar features.

3 MULTIPLE CLASSIFIERS

Over the last two decades, many multiple classifiers have been proposed. Several classifiers are trained and their results combined through a voting process. In this paper, we focus on multiple classifiers that are built at data level. The modified data sets are applied to train each classifier in the ensemble, the base classifier should be unstable, that is, small changes in the training set will lead to large changes in the classifier output. Neural

Type	Symbol	Feature
Height features	Δz	Height diff. to the ground
	Δz_{fl}	Height diff. first-last echoes
	σ_z^2	Height variance
	$\mathcal{C}$	Local curvature
Echo features	N	Number of echoes
	N_e	Normalized number of echoes
Eigenvalue features	λ_1	Highest eigenvalue
	λ_2	Medium eigenvalue
	λ_3	Lowest eigenvalue
	A_λ	Anisotropy
	P_λ	Planarity
	S_λ	Sphericity
	L_λ	Linearity
Local plane features	N_z	Deviation angle
	$\sigma_{n_z}^2$	Variance of deviation angles
	$\mathcal{R}_z$	Residuals to the local plane
	D_Π	Distance to the local plane
FW Lidar features	A	Echo amplitude
	w	Echo width
	σ	Cross-section
	α	Echo shape

Table 1: Synthesis of Lidar features for classification

Networks and Decision Trees are two examples of unstable classifiers. The most widely used methods are bagging and boosting. Boosting iteratively reproduces the training set for each classifier by increasing the weight to the incorrectly classified samples in previous classifier. Bagging is the acronym of "bootstrap aggregating". It is made of the ensemble of bootstrap-inspired classifiers produced by sampling with replacement from training instances and uses these classifiers to get an aggregated classifier. It aims at reducing the variance of a classifier (Briem et al., 2002).

3.1 Random Forests

Random Forests are a variant of bagging proposed by (Breiman, 2001). It is a decision tree based ensemble classifier that has excellent performance in classification tasks comparable to boosting (Breiman, 2001), even Support Vector Machines (SVMs) (Pal, 2005). It can be used with multi-class problems. It is non-parametric and does not require assuptions on the distribution of the data. This is an interesting property when different types or scales of input attributes are used. Moreover, Random Forests run efficiently on large data sets and can handle thousands of input variables without variable deletion. They do not overfit. They have a good predictive performance even when most predictive variables are noisy. Therefore, there is no need for variable preselection (Strobl et al., 2007). In addition, Random Forests is a classification algorithm that directly provide measures of variable importance (related to the relevance of each variable in the classfication process). These outstanding features make it suitable for the classification of remote sensing data such as multispectral data (Pal, 2005) or multisource data (Gislason et al., 2006).

Random Forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001). As in Breiman's method, in training, the algorithm creates multiple bootstrapped samples of the original training data, then builds a number of no pruning Classification and Regression Trees (CART) from each bootstrapped samples set and only a randomly selected subset of the input variables is used to split each node of CART. For classification, each tree in the Random Forests gives a unit vote for the most popular class at each input instance. The label of input instance is determined by a majority vote of the trees. The number of variables M, randomly chosen at each split, is considered as the single user-defined adjustable parameter. This parameter is not critical and is often set to the square root of the number of inputs (Gislason et al., 2006).

In this work, Random Forests are applied to classify airborne lidar data on urban scenes and to select the most important features for this task. This avoids the user to manually select relevant attributes. Random Forests have not yet been used with airborne lidar data.

3.2 Variable importance

Aside from classification, Random Forests provide measures of variable importance based on the permutation importance measure which was shown to be a reliable indicator (Strobl et al., 2007).When the training set for a particular tree is drawn by sampling with replacement, about one-third of the cases are left out of the sample set. These out-of-bag (OOB) data can be used to estimate the test accuracy and the permutation importance measure. The importance of variable m can be estimated by randomly permuting all the values of the mth variable in the OOB samples for each tree. A measure for variable importance (Breiman, 2001) can be the difference in prediction accuracy (i.e. the number of observations classified correctly) before and after permuting variable m, averaged over all trees. A high decrease of prediction accuracy indicates the importance of that variable.

3.3 Feature selection

The objective of feature selection is to identify small sets of lidar features that can still achieve a good predictive performance and so that correlated features should not be selected. In this paper, we use a backward elimination of features using OOB errors. This feature selection process was proposed in (Díaz-Uriarte and de Andrés, 2006), for biological application, to select genes of microarray data. Using data simulations, the authors showed its robustness to noise or redundant features.

To select the most relevant features, we iteratively fit Random Forests. At each iteration, a fraction of the features (the least important ones) is eliminated and a new forest is built. By default, the fraction is fixed to 0.2. It allows a relatively fast operation, and increases the resolution as the number of considered features becomes smaller. After fitting all forests, the selected set of features is the one whose OOB error rate is within u = 1 standard error of the minimum error rate of all forests. This is similar to the "s.e. rule" commonly used in classification trees literature (Breiman et al., 1984). This strategy can lead to solutions with fewer features while achieving an error rate that is not different, within sampling error, from the best solution.

4 EXPERIMENTAL RESULTS

4.1 Data set

The data acquisition was carried out with the RIEGL LMS-Q560 system over the city of Biberach (Germany). The main technical characteristics of this sensor are presented in (Wagner et al., 2006) and summarized in Table 2.

The city of Biberach includes artificial grounds, natural grounds, vegetation and buildings. This lidar data set has been used for classifying urban scenes using Support Vector Machines (Mallet

Flight height	Footprint size	PRF	Pulse density
500 m	0.25 m	100 KHz	2.5 (pt/m^2)

Table 2: Specifications of Biberach data set.

et al., 2008) with fewer features. The number of available reference samples is 797364 and they are split almost evenly between training and test samples.

Class	Training samples	Test samples
Buildings	187673	188015
Vegetation	15982	15723
Nat. Ground	2174	2149
Art. Ground	192704	192944
Total samples	398533	398831

Table 3: Data Set.

4.2 Variable importance results

The variable importance estimate for the training data is depicted on Figure 2 for each group feature. Considering all classes, it is obtained by the mean decrease permutation accuracy. The

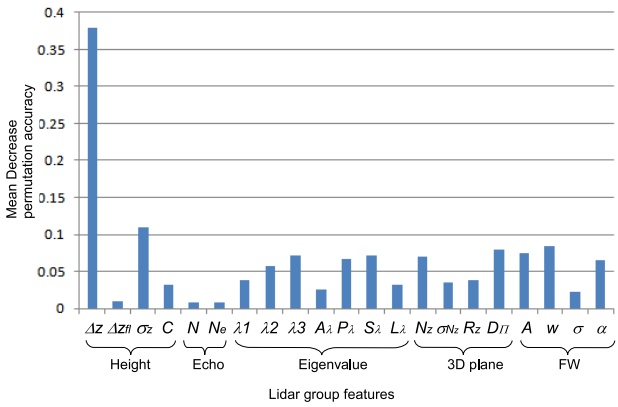


Figure 2: Variable importance of Lidar data by mean decrease permutation accuracy.

most important features are height-based: the height difference and the height variance. Echo-based features are not important for urban scene classification when using other attributes that describe more accurately the local distribution of 3D points such as eigenvalue-based or 3D plane-based features. Moreover, the First-Last height difference is not important to classify urban objects since it is used with the height variance σ_z^2. In fact, both variables can be correlated and the latter has more values which allows to distinguish rooftops building and ground for instance.

As for the eigenvalue-based features, λ_3 is the most important eigenvalue, as expected. In fact, it returns the lowest values for planar objects. The sphericity S_λ shows a high importance whereas the correlated anisotropy feature A_λ shows a lower one. This illustrates the advantage of permutation accuracy measure since redundant features should be less important. Among 3D-plane based features, the distance to plane D_Π seems to be the most important one. Finally, for full-waveform features, echo amplitude, and width are the most important for all classes. The FW cross-section σ is less important as it is correlated to the former features.

4.3 Feature selection results

To fix the Random Forests parameters, (Díaz-Uriarte and de Andrés, 2006) showed that the relation of OOB rate with the number of split variables M is largely independent of the number of trees T (for T between 1000-40000). In addition, the default setting of M is a good choice of OOB rate. Therefore, the feature selection was run with $M = 4$ and $T = 1000$ trees. Figure 3 is obtained by the backward iterative elimination of features using OOB errors. The graph is shown in a forward way to illustrate the more relevant features. On the x-axis, eigenvalue and FW Lidar features appear respectively in blue and red colors. The selection

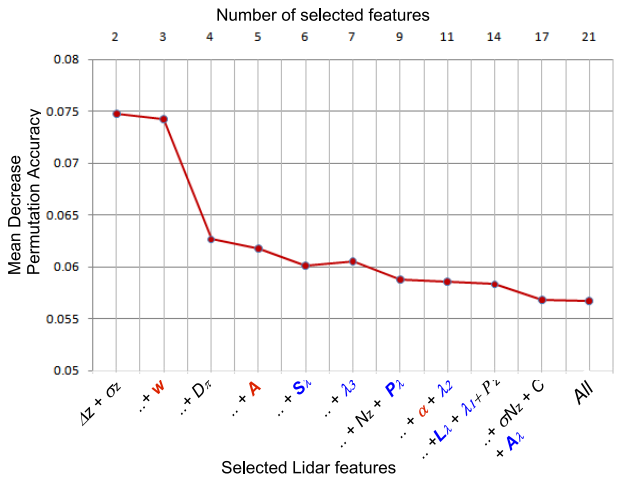


Figure 3: Iterative elimination Lidar feature selection.

process returns a feature vector of 17 attributes, where $\Delta_{z_{fl}}$, N, N_e and σ are eliminated. Another strategy may consist in keeping the set which first makes the variable importance decrease. In our case the final set may corresponds [Δ_z,σ_z,w,D_Π,A,S_λ]. The corresponding total error is 6%. One can observe that four feature groups are represented: the height based group is the most important one, then two FW features are selected which confirms the contribution of full-waveform lidar data for urban scene classification.

4.4 Classification results

Based on the 17 selected features, the Random Forests classification was run and variable importance computed for each class. Underlying parameters have been fixed to $M = 4$ which means that four variables are considered at each split and the number of trees was set experimentally to 60. The study area is visible on figure 4. We observe that errors occur mainly on building edges. The corresponding lidar points might be ambiguous since they correspond to transition points between building and artificial ground classes. Besides, these confusion errors are amplified due to the interpolation process of lidar points in 2D geometry.

The confusion matrix for test data is given in Table 4. The raster neighborhood size is 3×3. Artificial ground and buildings are well classified with lower error rate. However, the algorithm has more difficulties in classifying natural ground and vegetation. The former class suffers from smaller training set than the other classes. As for vegetation, confusions essentially occur with artificial ground due to the lidar data interpolation. In fact, in non-dense vegetated areas, the lidar beam is likely to reach the ground underneath and the resulting waveform has mixed properties. Therefore, some lidar feature pixels may be a combination of both classes.

4.4.1 2D Geometry impact Since lidar features are processed in a 2D geometry, we studied the effect of the neighborhood size on classification and variable importance results. A neighborhood size of 5×5 was tested. The corresponding confusion matrix is illustrated in Table 5.

Class	Art.Grnd	Build.	Nat.Grnd	Veget.	Error %
Art.Grnd	188562	3325	5	1052	2.3
Build.	13946	173545	5	519	7.7
Nat.Grnd	500	20	1622	7	24.5
Veget.	2604	566	0	12553	20.2

Table 4: Confusion matrix for test data using 60 trees and 4 split variables. Total error rate=5.65%, 3×3 window size

Class	Art.Grnd	Build.	Nat.Grnd	Veget.	Error %
Art.Grnd	186047	5483	33	1381	3.6
Build.	7364	180154	2	495	4.2
Nat.Grnd	1403	33	703	10	67.3
Veget.	3020	712	2	11989	31.1

Table 5: Confusion matrix for test data using 60 trees and 4 split variables. Total error rate=4.99%, 5×5 window size

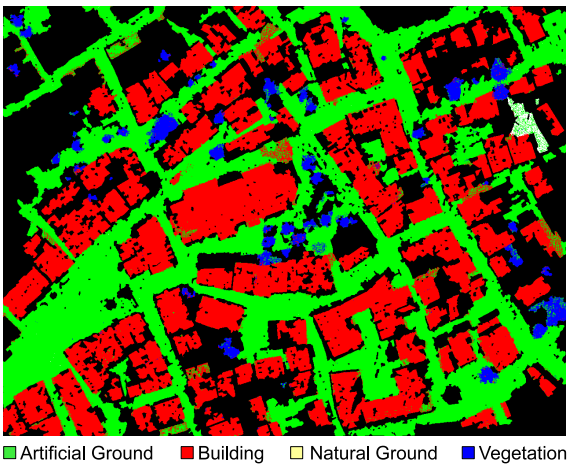
Figure 4 shows the corresponding classification result.

When increasing the neighborhood size, classification results are only enhanced for the building class, but are worse for all the other classes. In fact, buildings can be represented by large homogeneous segments with regard to lidar features. Therefore when dealing with larger neighborhoods, the training pixels have more representative mean values. It is similar for the artificial ground class however due to some small roads between buildings, the large neighborhood may include building points which increases the number of artificial ground pixels that are misclassified to building class (*cf.* Table 5). The total error rate seems to be smaller with a large neighborhood, however this is due to the fact that building class has the higher number of pixels.

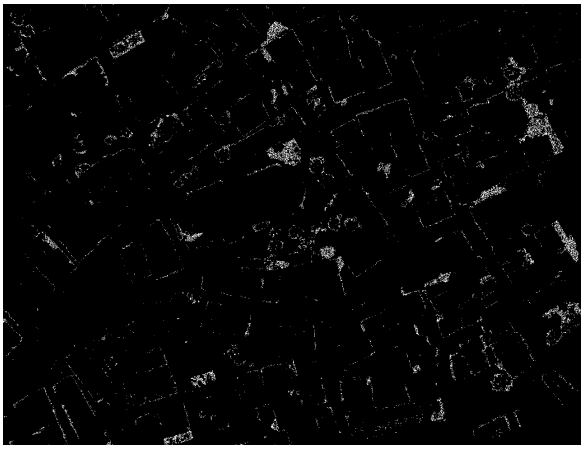
4.4.2 Variable Importance of selected features The variable importance was reprocessed for the 17 selected features in order to study more precisely the relevant features for each class. Two data sets were used with different window sizes: 3×3 (*cf.* Figure 5) and 5×5 (*cf.* Figure 6).

Firstly, one can observe that the variable importance values are higher and better distinguished when using a small window size since training pixels are more homogeneous. Secondly, when using a large window size, pixels may be mixed in the neighborhood, many variables give the same importance (*cf.* Figure. 6) for all classes which reveals the classification ambiguity. Consequently, the variable importances seem to be more reliable with a small neighborhood size. It will be discussed hereby.

We confirm that the height difference is the most important feature for all classes, which is the only topographic variable. For artificial ground, considering the different group features, the most important attributes are: P_λ for planarity, D_Π since it is a flat surface. FW selected attributes seem to have the same importance for this class. As for building class, the most important features in different groups are: S_λ,N_z and A. For plane-based features, the distance to the plane D_Π is more sensitive to the scan trajectory (the plane is better estimated along the scan trajectory since there are more points) and to the roof area (a large roof returns a higher distance to plane). For this reason it is less important than the deviation angle. For natural grounds, many variables seem to be important. This result should be interpreted with caution due to the few number of corresponding training samples. Finally, for vegetation pixels, the most relevant features of each group are: S_λ as it returns isotropic local distribution, R_z since no robust plane can be fitted to vegetation and finally the echo width w as already stated in (Wagner et al., 2008). For the latter classes, variable importance seems to be more dispersed between different attributes. This dispersion is correlated to the higher error rates on both classes (*cf.* Table 4).



(a) Classification results - Lidar Features : 5×5 window size



(b) Difference with the ground truth

Figure 4: Classification result (T=60 trees and M=4.)

4.4.3 Computing time The computing time for feature selection process (21 features) to 17 selected lidar features is 367 minutes. This value is not critical since the feature selection is done only one time provided a set of input attributes. The classification process for 17 lidar selected features needs 11.36 minutes. Random Forests do not need an intensive computing time, however a considerable amount of memory is needed to store a N by T matrix in memory.

5 CONCLUSION

In this work, Random Forests were successfully applied to the lidar feature selection to classify urban scenes. 21 lidar features were proposed and separated into five groups: height-based, echo-based, eigenvalue-based, plane-based and FW features. The method is a decision tree based ensemble classifier. It provides accurate classification and a variable importance estimate based on the permutation accuracy criteria. A feature selection was processed by iterative backward feature elimination. A minimal feature vector with 6 features provides a low OOB error rate. The most relevant feature for all classes is the height difference. Echo-based attributes seem to be non-relevant.Two FW features A and W appear in the final set, which confirms the contribution of full-waveform lidar for urban scene classification. The 2D window size impact on variable importance has been studied. Small sizes should be used to enhance the feature discrimination and to improve classification accuracy. Some non-relevant lidar features in our context, should be more useful for a finer classification such

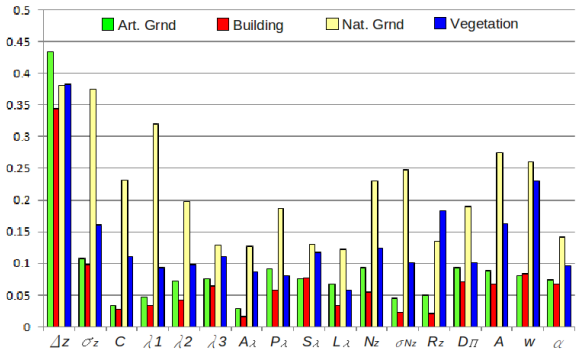


Figure 5: Variable importance per class - 3×3 window size

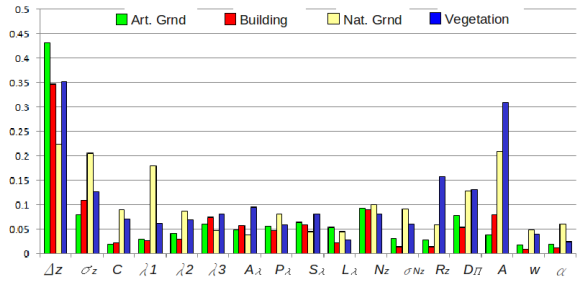


Figure 6: Variable importance per class - 5×5 window size

as the deviation angles for roof segmentation.

ACKNOWLEDGEMENTS

The Random Forests implementation software by L. Breiman and A. Cutler was used in experiments. For this work, we used the R interface of the software which is distributed freely on `http://www.r-project.org`.

REFERENCES

Breiman, L., 2001. Random forests. Machine Learning 45(1), pp. 5–32.

Breiman, L., Friedman, J., Olshen, R. and Stone, C., 1984. Classification and regression trees. Chapman & Hall, New York.

Briem, G., Benediktsson, J. and Sveinsson, J., 2002. Multiple classifiers applied to multisource remote sensing data. IEEE Transactions on Geoscience and Remote Sensing 40(10), pp. 2291–2999.

Charaniya, A., Manduchi, R. and Lodha, S., 2004. Supervised parametric classification of aerial lidar data. In: Real-Time 3D Sensors and their Use Workshop, in conjunction with IEEE CVPR, pp. 30–37.

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M. and Puech, W., 2007. Processing full-waveform lidar data: modelling raw signals. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 Part 3/W52, pp. 102–107. Espoo, Finland.

Díaz-Uriarte, R. and de Andrés, S. A., 2006. Gene selection and classification of microarray data using random forest. BMC BioInformatics, 7(3).

Ducic, V., Hollaus, M., Ullrich, A., Wagner, W. and Melzer, T., 2006. 3D Vegetation mapping and classification using full-waveform laser scanning. In: EARSeL and ISPRS Workshop on 3D Remote Sensing in Forestry, Vienna, Austria, pp. 211–217.

Gislason, P., Benediktsson, J. and Sveinsson, J., 2006. Random forests for land cover classification. Pattern Recognition Letters 27(4), pp. 294–300.

Gross, H. and Thoennessen, U., 2006. Extraction of lines from laser point clouds. In: ISPRS Conference Photogrammetric Image Analysis (PIA), Vol. 36 Part 3A, IAPRS, Bonn, Germany, pp. 87–91.

Gross, H., Jutzi, B. and Thoennessen, U., 2007. Segmentation of tree regions using data of a full-waveform laser. In: ISPRS Conference Photogrammetric Image Analysis (PIA), Vol. 36, IAPRS, Munich, Germany.

Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry & Remote Sensing 54(2-3), pp. 130–137.

Mallet, C., Bretar, F. and Soergel, U., 2008. Analysis of full-waveform lidar data for classification of urban areas. Photogrammetrie Fernerkundung GeoInformation (PFG) 5, pp. 337–349.

Matikainen, L., Kaartinen, H. and Hyyppä, J., 2007. Classification tree based building detection from laser scanner and aerial image datas. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), pp. 280–287.

Pal, M., 2005. Random Forest classifier for remote sensing classification. International Journal of Remote Sensing 26(1), pp. 217–222.

Rottensteiner, F., Trinder, J., Clode, S. and Kubik, K., 2005. Using the dempster-shafer method for the fusion of lidar data and multi-spectral images for building detection. Information Fusion 6(4), pp. 283–300.

Secord, J. and Zakhor, A., 2007. Tree detection in urban regions using aerial lidar and image data. IEEE Geoscience and Remote Sensing Letters 4(2), pp. 196–200.

Sithole, G. and Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing 59(1-2), pp. 85–101.

Steinle, E. and Vögtle, T., 2001. Automated extraction and reconstruction of buildings in laserscanning data for disaster management. In: Proc. of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona, Switzerland.

Strobl, C., Boulesteix, A.-L. and Augustin, T., 2007. Unbiased split selection for classification trees based on the Gini index. Computational Statistics & Data Analysis 52(1), pp. 483–501.

Wagner, W., Hollaus, M., Briese, C. and Ducic, V., 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. International Journal of Remote Sensing 29(5), pp. 1433–1452.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T. and Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry & Remote Sensing 60(2), pp. 100–112.

Xu, G. and Zhang, Z., 1996. Epipolar Geometry in stereo, motion and object recognition. Kluwer Academic Publishers.

Approche en 3D

Author's personal copy

ISPRS Journal of Photogrammetry and Remote Sensing 66 (2011) 56–66

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests

Li Guo [a], Nesrine Chehata [a,b,*], Clément Mallet [b], Samia Boukir [a]

[a] Institut EGID, Université de Bordeaux, Laboratoire GHYMAC 1 allée F. Daguin 33670 Pessac, France

[b] Université Paris Est, IGN, Laboratoire MATIS 73 avenue de Paris 94165 Saint-Mandé, France

ARTICLE INFO

Article history:
Received 8 April 2010
Received in revised form
19 August 2010
Accepted 20 August 2010
Available online 22 September 2010

Keywords:
Lidar
Multispectral image
Urban
Random forests
Variable importance

ABSTRACT

Airborne lidar systems have become a source for the acquisition of elevation data. They provide georeferenced, irregularly distributed 3D point clouds of high altimetric accuracy. Moreover, these systems can provide for a single laser pulse, multiple returns or echoes, which correspond to different illuminated objects. In addition to multi-echo laser scanners, full-waveform systems are able to record 1D signals representing a train of echoes caused by reflections at different targets. These systems provide more information about the structure and the physical characteristics of the targets. Many approaches have been developed, for urban mapping, based on aerial lidar solely or combined with multispectral image data. However, they have not assessed the importance of input features. In this paper, we focus on a multi-source framework using aerial lidar (multi-echo and full waveform) and aerial multispectral image data. We aim to study the feature relevance for dense urban scenes. The Random Forests algorithm is chosen as a classifier: it runs efficiently on large datasets, and provides measures of feature importance for each class. The margin theory is used as a confidence measure of the classifier, and to confirm the relevance of input features for urban classification. The quantitative results confirm the importance of the joint use of optical multispectral and lidar data. Moreover, the relevance of full-waveform lidar features is demonstrated for building and vegetation area discrimination.

© 2010 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

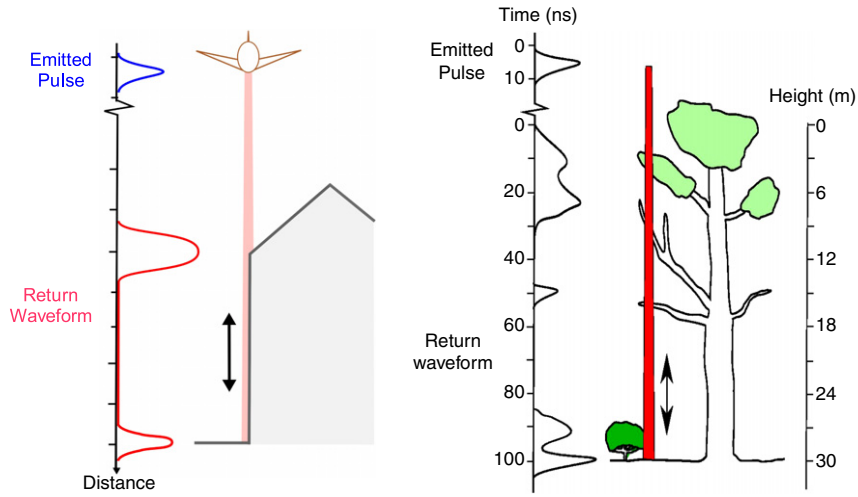
Airborne lidar systems have become a source for the acquisition of altimeter data. The development of various approaches based on lidar data for urban mapping has been an important issue for the last few years (Brenner, 2010). Many authors have shown the potential of multi-echo lidar data for urban area analysis and building extraction based on filtering and segmentation processes (Matikainen et al., 2003; Sithole and Vosselman, 2004; Rottensteiner et al., 2007; Matei et al., 2008). Classification is used for urban mapping to focus on building or vegetation areas, before the modeling step (Haala and Brenner, 1999; Poullis and Yu, 2009; Zhou and Neumann, 2009). Several classification methods were applied to lidar data for urban scenes such as the unsupervised Mean-Shift algorithm (Melzer, 2007), supervised classification such as Support Vector Machines (SVM) (Secord and Zakhor, 2007; Mallet et al., 2008), and a cascade of binary classifiers based on 3D shape analysis (Carlberg et al., 2009). Lidar classification can be based on geometric and textural features (Matikainen et al., 2003). Other methods include the lidar intensity (Charaniya et al., 2004), and combined lidar and multispectral data (Rottensteiner et al., 2007; Secord and Zakhor, 2007).

Recently, the full-waveform (FW) lidar technology (Mallet and Bretar, 2009) has emerged with the ability to record 1D signals representing multiple modes (echoes) caused by reflections at different targets (*cf.* Fig. 1). Thus, in addition to range measurements, further physical properties of targets may be revealed by waveform processing and echo fitting. In urban scenes, the potential of such data has been essentially investigated for urban vegetation with high density point clouds. In Gross et al. (2007) and Wagner et al. (2008), geometric and FW lidar features are derived from a FW 3D point cloud and used jointly to discriminate vegetated areas. In Rutzinger et al. (2008), the authors present an object-based analysis of FW lidar point cloud to extract urban vegetation. The 3D point cloud is first slightly over-segmented using a seeded region growing algorithm based on the echo width. Each segment is characterized by basic statistics (minimum, maximum, standard deviation, etc.) computed for the selected point features: amplitude, echo width and geometrical attributes. A supervised classification per statistical tree decision is then applied. In Höfle and Hollaus (2010), an improved echo ratio feature is computed

* Corresponding author at: Institut EGID, Université de Bordeaux, Laboratoire GHYMAC 1 allée F. Daguin 33670 Pessac, France. Tel.: +33 662007836.
E-mail addresses: nesrine.chehata@egid.u-bordeaux3.fr, nesrine.chehata@ign.fr (N. Chehata).

0924-2716/$ – see front matter © 2010 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2010.08.007

Author's personal copy



(a) Building area. (b) Vegetation area.

Fig. 1. Transmitted and received signals with a small-footprint FW lidar.

allowing a better vegetation discrimination. A rule-based classification is refined using FW echo amplitude and width to discriminate non-vegetation objects such as building walls, roof edges, and powerlines. Mallet et al. (2008) studied the contribution of FW lidar data for urban scene classification using 3D features and SVM classification.

All these works focus on the classification process but do not analyze the relevance of input features. The contribution of lidar data in comparison with multispectral images is not quantified, nor is the importance of FW lidar features for urban scenes. Our objective is to study the relevance of multi-source data composed of lidar features (multi-echo (ME) and full waveform (FW)) and multispectral *RGB* features for mapping urban scenes. Four urban classes are considered: buildings, vegetation, artificial ground, and natural ground. Artificial ground gathers all kinds of streets and street items such as cars and traffic lights whereas natural ground includes grass, sand and bare-earth regions. To achieve this goal, the Random Forests classifier is chosen. This algorithm is well suited to a multi-source framework and is able to process large datasets. Besides, it measures feature importance so that its contribution can be examined for the different classes we consider. In our previous work Chehata et al. (2009b), feature importance was studied for 2D ME and FW lidar features and multispectral data. In Chehata et al. (2009a), 21 lidar features were generated and analyzed, and the influence of the 2D window size was studied. In this work, we present a global framework to study the relevance of features for classification using the Random Forests classifier. Besides, to confirm these observations, the margin theory is used as a confidence measure of the classifier, which consequently helps in evaluating the suitability of input features. The methodology is applied in this case to lidar and multispectral data. The remainder of this paper is organized as follows. The lidar and optical features are introduced in Section 2. In Section 3, Random Forests classification, the feature importance measure, and a new margin are presented. Experimental results are then given and discussed in Section 5. Finally, conclusions and perspectives are drawn.

2. Multi-source airborne lidar and image features

The multi-source data is composed of an orthoimage and a full-waveform lidar dataset. The orthoimage is composed of three multispectral bands in the visible domain: Red, Green and Blue. Lidar and image data are complementary: the optical image provides high spatial resolution and multispectral reflectances in the visible domain, while lidar uses the infrared domain and provides 3D geometric information, but the data is under-sampled. Moreover, it has the ability to penetrate the vegetation, giving more information about these areas (Brenner, 2010). Finally, FW lidar provides more information about the physical properties of the targets. The data are georeferenced and processed in the same geographic projection system. In order to combine data from sources with distinct geometries, lidar points are projected onto a regular 2D raster map. One image is generated per feature used for classification. Red (R), Green (G) and Blue (B) channels of the orthoimage are used as three independent optical features.

The feature vector is composed of twelve features: three optical features R, G and B, five multi-echo lidar features and four full-waveform lidar features. Lidar features are obtained by lidar waveform modeling. It consists in decomposing the waveform into a sum of components or echoes, so as to characterize the different targets along the path of the laser beam. A parametric approach is chosen. Parameters of an analytical function are estimated for each detected peak in the signal. Generally, the received signal is decomposed by fitting Gaussians to the waveform (Wagner et al., 2006; Reitberger et al., 2009). The waveform fitting is processed by an iterative Levenberg–Marquardt technique. This, first, leads to multiple echo detection and range measurements. In our study, ME lidar features are derived from a FW lidar dataset providing a higher altimetric accuracy. However, they can be obtained directly using a multi-echo lidar system.

For each pixel, the lidar features are computed using the 3D points included in a given cylindrical neighborhood ν_P, centered at the current pixel P and defined by the parameter r (*cf.* Fig. 2(a)).

The raster cell spacing c and the cylinder radius r are chosen with respect to the following:

- the 3D point density: a minimal number of lidar points is necessary to compute an unbiased local plane Π_P;
- the contrast between objects that we aim to retrieve: a small value of c combined with a high value of r can lead, with a dense point cloud, to smoothed images, and to mix in a single pixel, different kinds of objects (see Fig. 2(a)). In such cases, the geometric spatial features may be affected, thus providing biased values.

We assume that a 3D neighborhood should include at least five points to process lidar features. The minimal radius is fixed in concordance with the 3D point density. The maximal radius equals

Author's personal copy

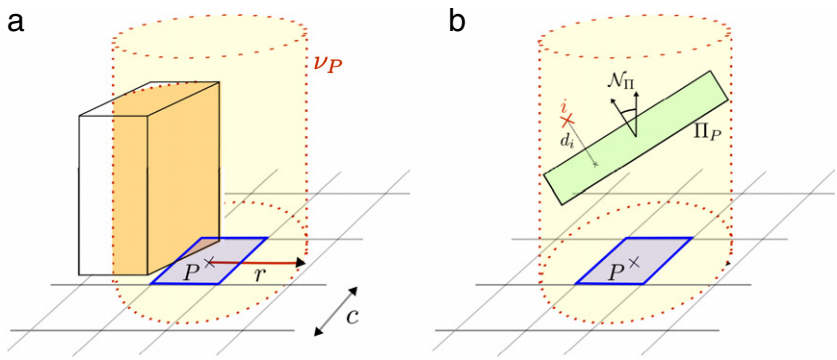


Fig. 2. 3D neighborhood for lidar feature computation.

the maximal object size in the image. For each 3D neighborhood, the radius is increased iteratively until validating our assumption. Values of c and r for our dataset are given in Section 4.

The five multi-echo lidar features are spatial. In urban scenes, most objects can be described by planar surfaces such as building roofs and roads. The planarity of the local neighborhood should help in discriminating buildings from vegetation. The local plane Π_P is estimated using a robust M-estimator with norm $L_{1.2}$ (Xu and Zhang, 1996) on the points included in ν_P. Such a norm provides a plane estimator scarcely affected by outlying data, such as superstructures for buildings, and low-rise objects for ground areas. Several features are derived from the computation of Π_P.

- Δz: height difference between the current lidar point P and the lowest point found in a large cylindrical volume whose radius r depends on the size of the largest structure in the area of interest. It has been experimentally set to 15 m. This feature will help in discriminating between ground and off-ground objects without a digital terrain model (DTM) estimation.
- N_z: deviation angle of the normal vector of the local plane Π_P from the vertical direction (see Fig. 2(b)). This feature highlights the ground with the lowest values.
- $\mathcal{R}_z$: residuals of the local plane estimated within the small radius cylinder. Residuals $\mathcal{R}_z$ are calculated with respect to the estimated plane as follows:

$$\mathcal{R}_z = \sum_{i \in \nu_P} \frac{(d_i)^l}{l} \tag{1}$$

where d_i is the distance between the lidar point $P_i \in \nu_P$ (ν_P is the local 3D neighborhood) and the plane Π_P (see Fig. 2(b)). L_l is the norm used for estimating the 3D plane. Here $l = 1.2$. Residuals should be high for vegetation.

A single emitted laser pulse may generate several echoes from objects located at different positions inside the pulse conical 3D volume. In urban scenes, this is particularly interesting for vegetated areas and building edges since several echoes will be recorded per emitted pulse. Consequently, features based on the number of echoes should be relevant for urban scene classification. Two features have therefore been selected.

- N: total number of echoes within each waveform of the current pixel P. This feature will be prominent for vegetation and building facades (*cf.* Fig. 3(e)).
- N_e: normalized number of echoes obtained by dividing the echo number (position of the echo within the waveform) by the total number of echoes within the waveforms of P. This feature highlights the vegetation since multiple reflections can occur on it (*cf.* Fig. 3(f)).

The remaining features are more specific to FW lidar data. The received signal is generally decomposed by fitting Gaussians to the waveform. However, in urban areas, the characteristics of peaks may differ significantly due to multiple geometric and radiometric effects of the targets (*e.g.*, roof slopes and various materials). Consequently, other modeling functions have been proposed. Chauve et al. (2007) have improved the signal fitting using the generalized Gaussian function. We used the latter modeling to retrieve the following FW lidar features.

- A (echo amplitude): high amplitude values can be found on building roofs, on gravel, and on cars. Asphalt and tar streets have low values. The lowest values correspond to vegetation due to the penetration of the laser pulse and the signal attenuation (*cf.* Fig. 3(g)).
- w (echo width): higher values correspond to vegetation since it spreads the lidar pulses. A narrow pulse is likely to correspond to ground and buildings. However, the width value increases with the slope (*cf.* Fig. 3(h)).
- σ (echo cross-section (Wagner et al., 2008)): the values are high for buildings, medium for vegetation, and low for artificial ground (*cf.* Fig. 3(i)).
- α: echo shape, describing how locally flattened the echo is. Chauve et al. (2007) showed that very low and high shape values correspond respectively to building roofs and vegetation.

The feature vector f_v is defined as follows with three optical features R, G and B, five multi-echo lidar features and four full-waveform lidar features.

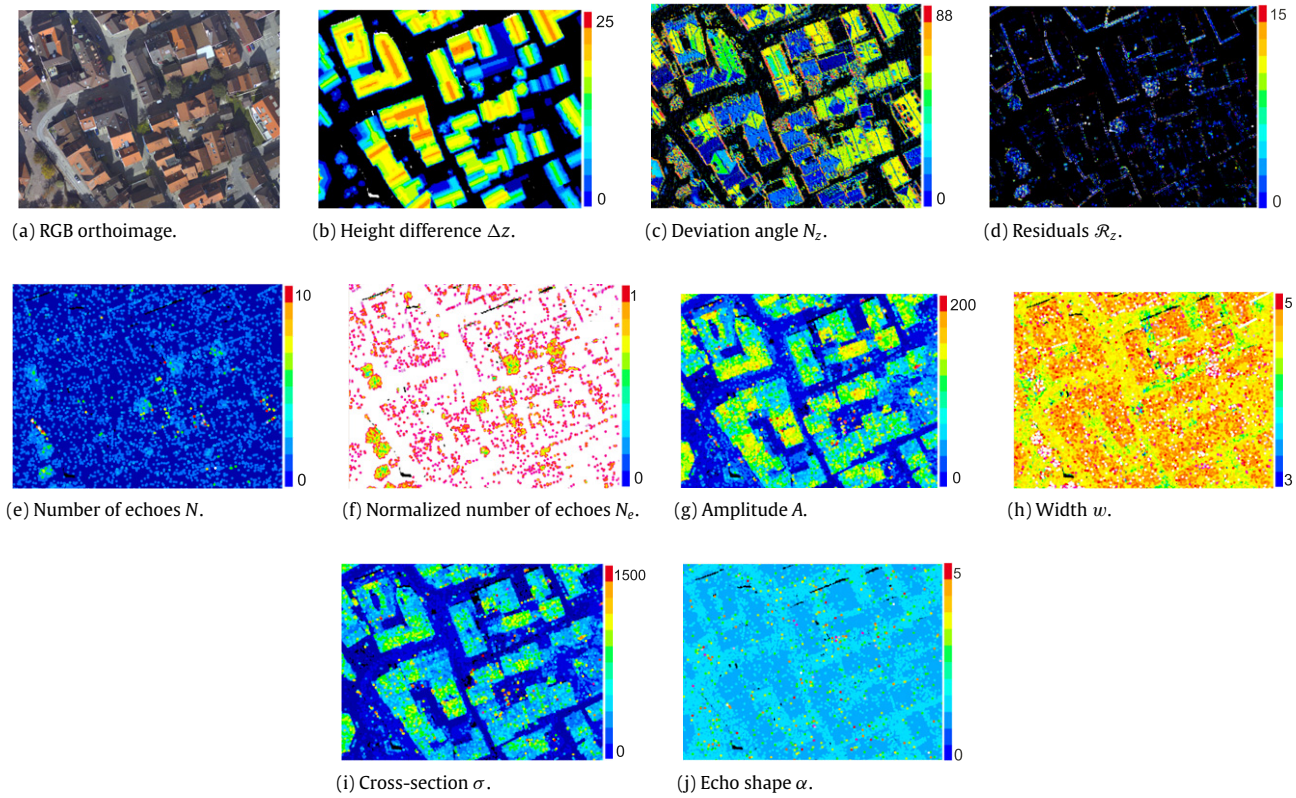
$$f_v = [R\ G\ B;\ \Delta z\ N_z\ \mathcal{R}_z\ N\ N_e;\ A\ w\ \sigma\ \alpha]^T. \tag{2}$$

Table 1 summarizes the expected lidar feature values for the different classes we consider on urban scenes.

Fig. 4 depicts the median feature values for each class. The values are processed from the training dataset. A logarithmic scale is used for more clarity on small values. Feature values that equal 0 do not appear on such graphics. This is the case, for instance for Δz, for both types of grounds, and R_z for all classes except vegetation.

At first glance, one can notice the most discriminative features for all classes such as R, B, Δz, N_z, A, σ. Features N_z, R_z, N_e better discriminate the vegetation class. Ground classes give generally the same values except for optical image bands, and for FW lidar features A and σ. The second analysis consists in observing the lidar features (Fig. 3). Visually, we can assess whether a feature is important for the classification or not. For instance, FW lidar features (A and σ, Fig. 3(g) and (i)) lead to homogeneous values for buildings, hence minimizing this class intra-variance. FW lidar features are consequently well suited to the building class. Conversely, the echo width (w) values (Fig. 3(h)) are very noisy and do not help in discriminating the four classes visually. Moreover, when comparing the number of echoes and the normalized

Author's personal copy



(a) RGB orthoimage. (b) Height difference Δz. (c) Deviation angle N_z. (d) Residuals $\mathcal{R}_z$.

(e) Number of echoes N. (f) Normalized number of echoes N_e. (g) Amplitude A. (h) Width w.

(i) Cross-section σ. (j) Echo shape α.

Fig. 3. 2D optical, multi-echo and full-waveform lidar features.

Table 1
Empirical values of lidar features for the four classes and class suitability.

Lidar	Feature	Building	Vegetation	Artificial ground	Natural ground	Suitability
ME	Δz	Variable	Variable	$\rightarrow 0$	$\rightarrow 0$	Ground vs off-ground
	N_z	$[0°, 45°]$	Variable	$[0°, 10°]$	$[0°, 10°]$	Ground
	$\mathcal{R}_z$	$\rightarrow 0$	High	$\rightarrow 0$	$\rightarrow 0$	Vegetation
	N	~ 1	>1	~ 1	1	Vegetation
	N_e	~ 1	≤ 1	~ 1	1	Vegetation
FW	A	Variable	Medium	Low	Variable	Artificial ground
	w	Medium	High	Variable	Variable	Vegetation
	σ	High	Medium	Medium	Variable	Building
	α	Variable	Variable	$\simeq\sqrt{2}$	$\simeq\sqrt{2}$	Natural ground

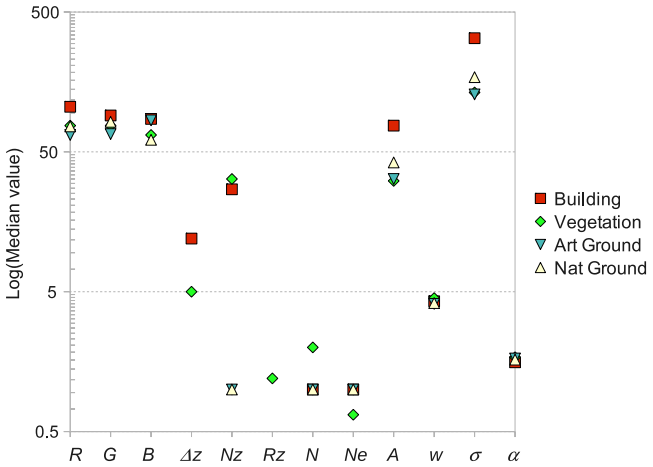


Fig. 4. Lidar features' median values per class.

number of echoes, the latter better highlights the vegetation. These observations will be confirmed by the feature importance measures in Section 5.1.

3. Random Forests classifier

Random Forests (RF) are a variant of bagging proposed by Breiman (Breiman, 2001). It is a decision-tree-based ensemble classifier that can achieve a classification accuracy comparable to boosting (Breiman, 2001), or SVM (Pal, 2005; Zhu, 2008). It does not overfit, runs fast and efficiently on large datasets such as lidar data. It does not require assumptions on the distribution of the data, which is interesting when different types or scales of input features are used. These outstanding properties make it suitable for remote sensing classification. It was successfully applied to multispectral data (Pal, 2005), multitemporal SAR images (Waske and Braun, 2009), hyperspectral data (Ham et al., 2005), or multi-source data (Gislason et al., 2006), where Landsat MSS and topographic data were used. Waske and Benediktsson (2007) applied it on SAR and multispectral images. For airborne multi-source classification using lidar and optical multispectral images, we showed in previous work that it achieves a classification accuracy comparable to SVM precision with a shorter training time (Chehata et al., 2009b).

In addition, the importance of each feature can be estimated during the training step. In this work, we exploit this property on

Author's personal copy

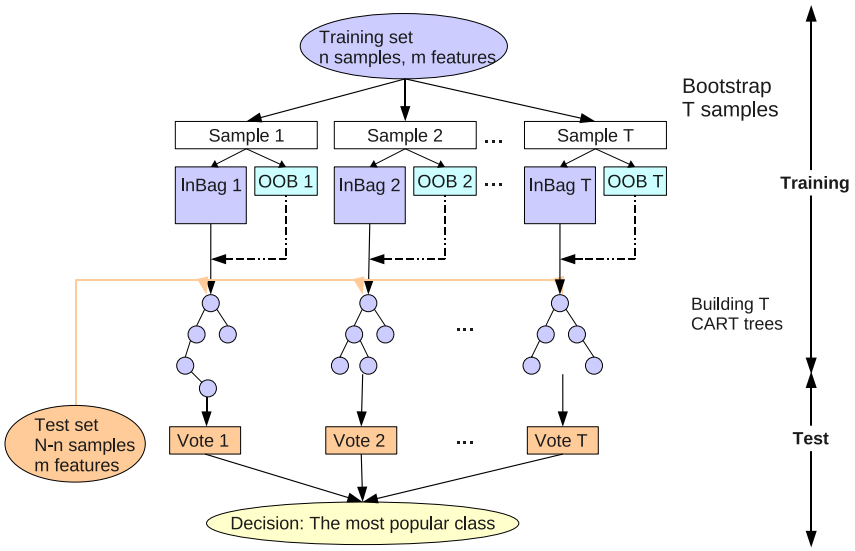


Fig. 5. The flow chart of Random Forests.

multi-source data in order to measure the relevance of airborne lidar and optical image features for classifying urban scenes.

3.1. Principle

Random Forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001). In training, the algorithm creates T multiple bootstrapped samples of the original training data, and then builds a number of no pruned Classification and Regression Trees (CART) from each bootstrapped sample set. Only a randomly selected subset of the input features is considered to split each node of CART. The feature that minimizes the Gini impurity is used for the split (Breiman, 2001). For classification, each tree gives a unit vote for the most popular class at each input instance. The final label is determined by a majority vote of all trees. The RF classifier has two parameters: the number of trees T and the number of variables M randomly chosen at each split. Breiman's Random Forests error rate depends on two parameters: the correlation between any pair of trees and the strength of each individual tree in the forest. Increasing the correlation increases the forest error rate while increasing the strength of the individual trees decreases this misclassification rate. Reducing M reduces both the correlation and the strength. M is often set to the square root of the number of inputs (Breiman, 2001).

A global scheme highlighting the main steps of Random Forests is depicted in Fig. 5.

When the training set for a particular tree is drawn by sampling with replacement, about one-third of the cases are left out of the sample set. These samples are called Out-of-Bag (OOB) data (*cf.* Fig. 5) and are used to estimate the feature importance as detailed hereby. In the following, we denote by $\mathcal{B}^{(t)}$ the In-Bag samples for a tree t and by $\mathcal{B}^{c(t)}$ the complementary samples, *i.e.*, the OOB data for the tree t.

3.2. Feature importance measure

Aside from classification, Random Forests provide a measure of feature importance that is processed on OOB data and is based on the permutation importance measure (Breiman, 2001). The importance of a feature f can be estimated by randomly permuting all the values of this feature in the OOB samples for each tree. This follows the idea that a random permutation of a feature mimics the absence of that feature from the model. The measure of feature importance is the difference between prediction accuracy (*i.e.*, the number of observations correctly classified) before and after permuting feature f, averaged over all the trees (Breiman, 2001). A high prediction accuracy decrease denotes the importance of that feature. Suppose that the training samples consist of pairs of the form (x_i, l_j) where x_i is an instance, and l_j its true label. The importance of a feature f per tree t is computed as follows:

$$FI^{(t)}(f) = \frac{\sum_{x_i \in \mathcal{B}^{c(t)}} I(l_j = c_i^{(t)})}{|\mathcal{B}^{c(t)}|} - \frac{\sum_{x_i \in \mathcal{B}^{c(t)}} I(l_j = c_{i,\pi_f}^{(t)})}{|\mathcal{B}^{c(t)}|} \tag{3}$$

where $\mathcal{B}^{c(t)}$ corresponds to OOB samples for a tree t, with $t \in \{1, \ldots, T\}$. $c_i^{(t)}$ and $c_{i,\pi_f}^{(t)}$ are the predicted classes for sample x_i before and after permuting the feature f respectively. Note that $FI^{(t)}(f) = 0$, if feature f is not in tree t. The importance score for a feature f is then computed as the mean importance over all trees:

$$FI(f) = \frac{\sum_T FI^{(t)}(f)}{T} \tag{4}$$

where T is the number of trees.

3.3. Margin definition

The margin concept of ensemble learning methods was first proposed by Schapire et al. (1998) to explain the success of boosting type algorithms. The concept was then generalized to analyze other types of ensemble classifiers (Breiman, 2001). Suppose that the training samples consist of pairs of the form (x_i, l_j), where x_i is an instance and l_j its true label. The margin m_i of instance x_i is computed as follows (Tang et al., 2006):

$$m_i = \text{margin}(x_i, l_j) = \frac{v_{i,l_j} - \sum_{c \neq l_j} v_{i,c}}{\sum_c v_{i,c}} \tag{5}$$

where v_{i,l_j} is the number of votes for the true class l_j, and $v_{i,c}$ is the number of votes for any class c with $c \neq l_j$. Hence, the margin is given by the difference between the fraction of classifiers voting correctly and incorrectly. It measures the strength of the vote. The margin ranges from -1 to $+1$. The positive margin value

Author's personal copy

Table 2
2D training and test samples. Class proportions show a highly imbalanced dataset.

Class	Training samples	Test samples	Proportion (%)
Building	17 617	71 398	46
Vegetation	1 616	6 752	4
Artificial ground	18 671	75 283	48
Natural ground	860	3 463	2
Total samples	38 764	156 896	100

of a sample indicates that this sample has been correctly classified, whereas a negative value means that the sample has been misclassified. The larger the margin, the more the confidence in the classification. A value close to 0 indicates a low confidence in the classification. Several studies have shown that the generalization performance of an ensemble classifier is related to the distribution of its margins on the training samples. Schapire et al. (1998) proved that achieving a larger margin on the training set leads to an improved bound on the generalization error of the ensemble.

In the following, the margin is exploited as a measure of confidence of the RF classifier. We use it to assess the relevance of input features.

4. Test dataset

The data is composed of georeferenced airborne lidar and multispectral *RGB* image. The lidar data acquisition was carried out with the RIEGL LMS-Q560 system over the city of Biberach (Germany). The main technical characteristics of this sensor are presented by Mallet and Bretar (2009). The lidar point cloud has a point density of approximatively 2.5 pts/m^2 with a footprint size of 0.25 m. The orthophotography has been captured with an Applanix DSS 22M device. Its resolution is 0.25 m and dimensions are 640 $*$ 485 pixels. To compute 2D lidar features, the following parameters were used: $c = 0.25$ m and $r_{\min} = 0.75$ m. This r value is chosen to provide the minimal number of lidar points (5 pts) in a 3D neighborhood.

The number of available reference samples is 195 660. 20% of randomly selected samples are used as a training set (*cf.* Table 2). One can observe that dense urban scenes are characterized by highly imbalanced classes: building and artificial ground are major classes while vegetation and natural ground are minor classes. The ground truth is processed manually, based on an oversegmentation of the orthoimage.

5. Results and discussion

The Random Forests implementation software by Leo Breiman and A. Cutler (http://www.r-project.org) was used in the experiments. Underlying parameters have been fixed to $M = \sqrt{n}$, where n is the number of features, and the number of trees T was set to 100.

Feature importance and margin confidence results will be presented in the first part, and then discussed with regard to physical lidar and multispectral image properties. The contribution of the considered features is detailed and explained not only for all classes, but also per class.

5.1. Feature importance results

To compute the feature importance, a balanced training set (3000 samples per class) is used to avoid biases due to a small number of samples of vegetation and natural ground classes in urban scenes. It is essential to select the best features for minor classes. A variable importance estimate for the training data is depicted in Fig. 6. The first three features are the optical

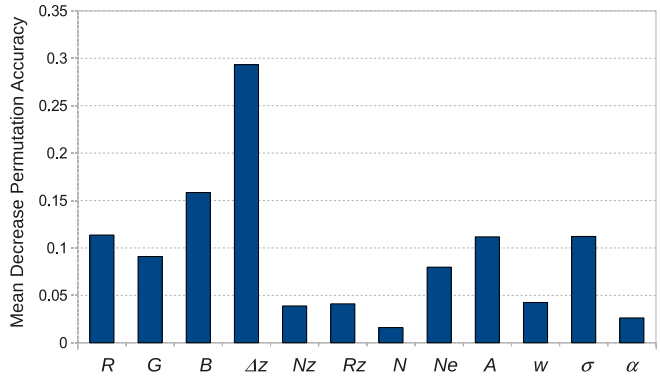


Fig. 6. Variable importance by mean decrease permutation accuracy. Balanced training data with 3000 samples per class.

components R, G, and B, whereas the latter are lidar features. The decrease permutation accuracy is averaged over all trees for all classes.

From Fig. 6, it appears that the most relevant features for all classes are the height difference Δz, red and blue channels R, B, the echo amplitude A, and the echo cross-section σ. This leads to the following optimal feature vector: $[\Delta z, R, B, A, \sigma]$ that includes both optical image and FW lidar features as shown by Chehata et al. (2009b). This preliminary result will be confirmed in the following sections based on the margin theory.

5.2. Margin results

The margin is a confidence measure of predictions which helps in evaluating the classification quality and the RF classifier strength. Fig. 7 depicts the study area, the ground truth and the training margin image. This image was produced in the training step. All data are used as training samples just to illustrate the training margins.

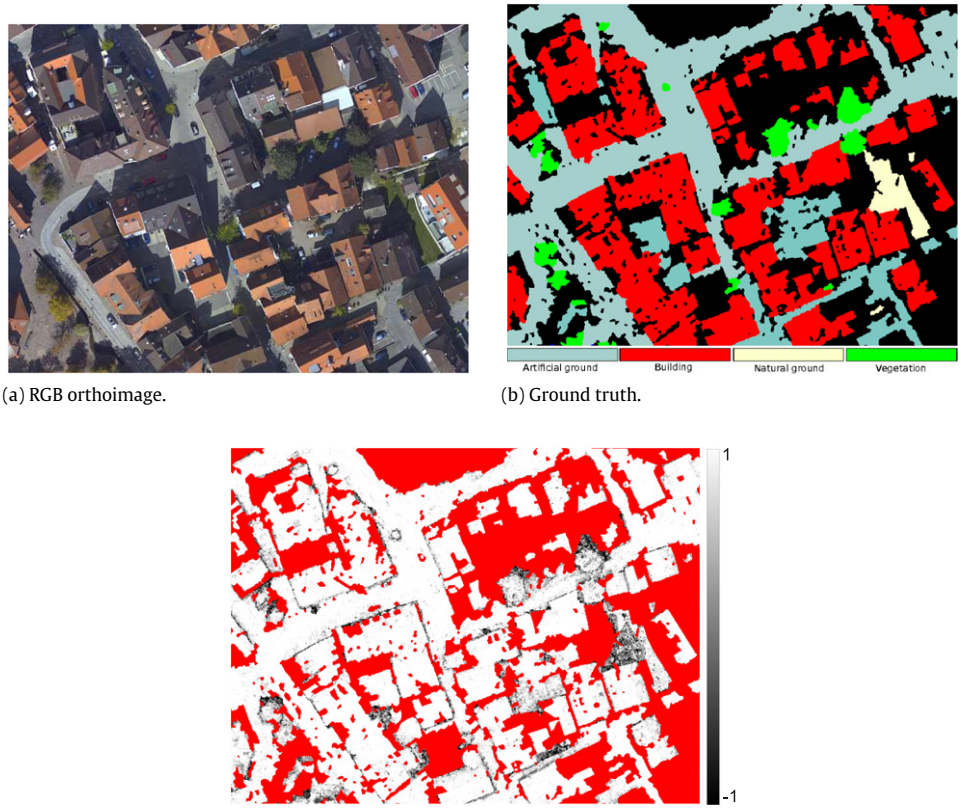
The margin values are significantly higher at the center of classes in the feature space whereas the smaller margins correspond mainly to the class boundaries or to noise. In our case, class boundaries are located on building facades which are a transition between building and artificial ground classes, and also on vegetation boundary pixels that mix artificial ground and vegetation information. In fact, laser pulses can penetrate vegetation and reach the ground underneath. Consequently, for these areas, pixels combine various classes which make them harder to classify, leading to a low margin value. Moreover, shadowed areas are likely to have low margins when using R and B channels which have uniform irrelevant values in these areas. This is the case of natural ground points on the right of Fig. 7, or in urban corridors. Finally, higher margin values are located, in 2D space, at the center of buildings and artificial grounds as these classes are well discriminated using the five selected features.

5.3. Margin and classification confidence

Fig. 8 illustrates the test data margin histograms for well classified and misclassified samples.

For well classified samples, 88% of samples have a good classification confidence (margin ≥ 0.7). For misclassified samples, the margin distribution is more scattered and the confidence varies. Indeed, there are several sources of errors such as noise samples, class boundaries or even ground-truth errors. Considering the extreme case for misclassified samples with a high confidence (margin $= -0.9$), this is probably due to errors in the ground truth.

Author's personal copy



(a) RGB orthoimage.

(b) Ground truth.

(c) Margin image; five best features.

Fig. 7. Margin image based on the five best features – Random Forests classification ($T = 100$, $M = 2$). Black regions on the ground-truth image are not labelled and they correspond to the red regions in the margin image.

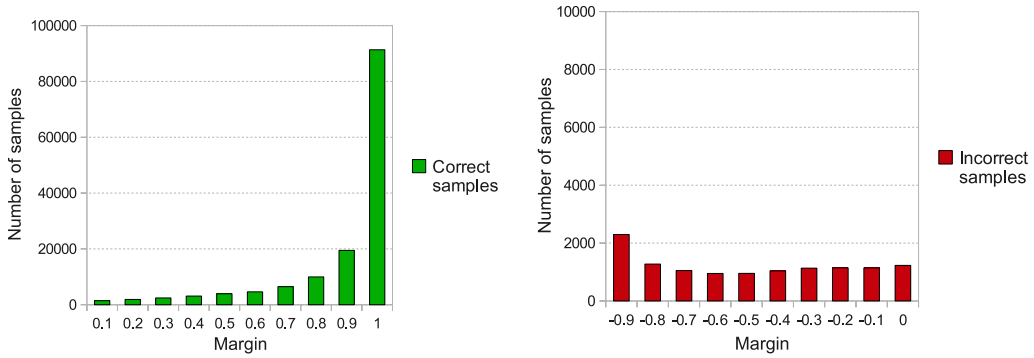


Fig. 8. Margin histograms for well classified and misclassified test data.

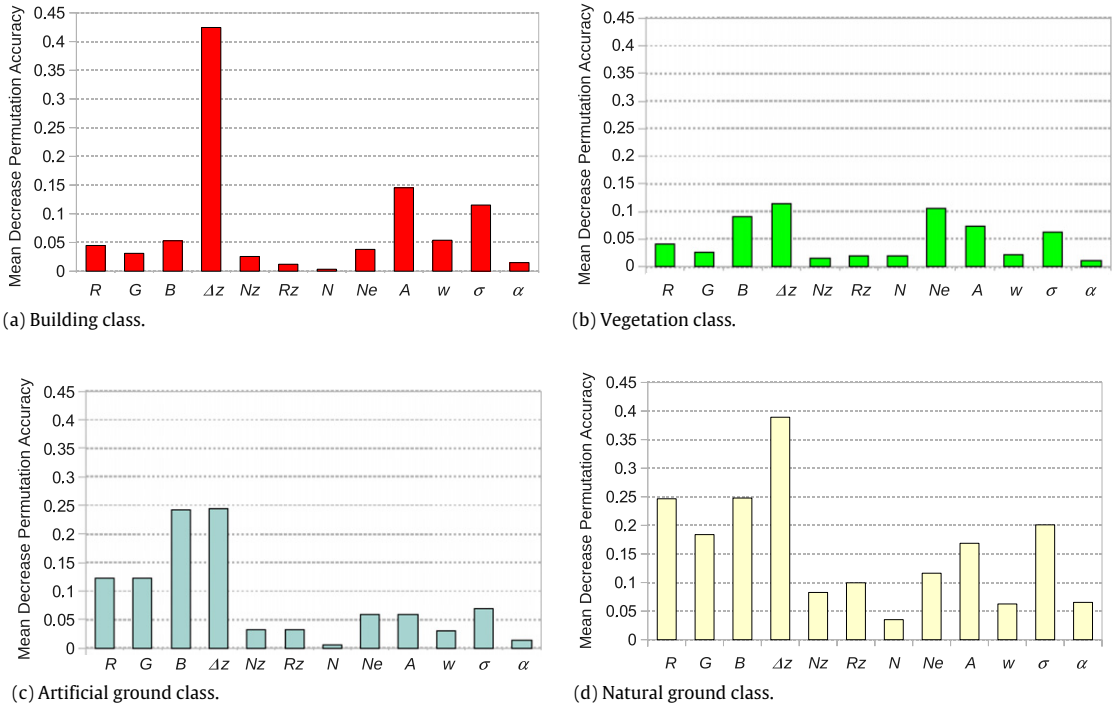
Table 3
Confusion matrix for test data using a RF classifier with the five best features, 100 trees and 2 split variables. Global error rate = 5.03%.

Obs.	Pred.				
	Building	Vegetation	Artificial ground	Natural ground	Omission error (%)
Building	69 128	242	1986	42	3.2
Vegetation	468	4882	1344	58	27.7
Artificial ground	1817	589	72 559	318	3.6
Natural ground	59	10	956	2438	29.6
Commission error (%)	3.3	14.7	5.7	14.6	5.0

5.4. Classification results

The Random Forests classifier is run with the five best selected features. The confusion matrix is given in Table 3. The training dataset is highly imbalanced with two major classes (building and artificial ground) that are more than ten times larger than vegetation and natural ground classes. We can notice that the artificial ground and buildings are well classified. However, the

Author's personal copy



(a) Building class.

(b) Vegetation class.

(c) Artificial ground class.

(d) Natural ground class.

Fig. 9. Feature importance per class using mean decrease permutation accuracy.

algorithm has more difficulties in classifying natural ground and vegetation which suffer from smaller training sets. Higher sources of errors are highlighted in Table 3.

Errors essentially occur between building and artificial ground since (1) these two classes have similar colors on optical images and the presence of shadows increases these errors, and (2) lidar features are ambiguous for building facades that are transitions between both classes. In addition, vegetation classes can be confused with artificial ground classes as the laser pulse can reach the ground under sparse vegetation. This confusion is increased by 2D lidar data interpolation. Finally, classification errors may occur between natural and artificial grounds since ME lidar features do not allow discriminating them.

Errors correspond to small margin values. They are depicted in Fig. 7.

5.5. Feature relevance for urban scenes

The feature relevance for urban scenes can be studied by the mean decrease permutation accuracy for all classes (*cf.* Fig. 6). From this figure, the height difference Δz appears to be the most important feature for all classes. It corresponds to the only topographic feature which helps in discriminating between ground and off-ground classes. Moreover, the study area is essentially composed of red brick rooftops and artificial ground which have high reflectances in the red and blue channels, respectively, leading to a high importance of both optical features. On the contrary, the number of echoes N is the less important feature. Indeed, this feature is not discriminating since there is one echo for all building rooftops and grounds, and it varies for vegetation. The other features will be discussed per class in the following section.

5.6. Feature relevance per class

Fig. 9 depicts the feature importance for each class. For the building class (Fig. 9(a)), the most important variables are Δz and two FW lidar features: A and σ. The latter features induce high values for buildings as explained in Section 2. Conversely, *RGB* channels are less important as they are not homogeneous on rooftops due to shadows as shown in Fig. 3(a). Both features that are related to the planarity of the local neighborhood (N_z, $\mathcal{R}_z$) show a low importance since they are very sensitive to the slope. These features are not homogeneous for the building class. However, they should be relevant for roof segmentation. As for the echo features, building facades lead to multiple echoes which explains the relative importance of the normalized number of echoes, N_e.

Regarding the vegetation class (Fig. 9(b)), in addition to the height difference, the normalized number of echoes N_e is very discriminating. A high number of echoes is specific to vegetation. Moreover, the normalized number of echoes is more important than the number of echoes since the values of the normalized number of echoes are bounded between 0 and 1 which is more appropriate for classification. Only one spectral band B, and FW lidar features A and σ are important. The latter features exhibit low values for vegetation as presented in Section 2.

As for multispectral bands, the spectral band G does not seem to be important in this case since it does not allow distinguishing vegetation from natural ground.

However, due to the laser properties, this class is composed of both vegetation and artificial ground information. Consequently, feature importances are more dispersed and height difference appears to be less important than for other classes.

The artificial ground class (Fig. 9(c)) shows more importance in multispectral *RGB* channels compared to lidar data. Indeed, due to urban corridors and building facades, points belonging to this class may also belong to the building class. Due to this confusion, height difference is less important.

As for the natural ground class (Fig. 9(d)), the feature importances are more dispersed between topographic, radiometric, and some lidar features which makes the classification more complex (*cf.* Table 3).

The echo shape feature seems to be more important for this class as low values of α correspond to the natural ground class (*cf.* Table 1). However, this feature is not much relevant for the

Author's personal copy

Table 4
Classification accuracy per class with respect to type and number of features.

Features	Building	Vegetation	Artificial ground	Natural ground	Total accuracy
RGB (3)	0.80	0.39	0.89	0.43	0.82
ME lidar (5)	0.96	0.50	0.95	0.01	0.92
FW lidar (4)	0.95	0.69	0.95	0.55	0.93
ME + FW lidar (9)	0.97	0.74	0.96	0.37	0.94
Selected features (5)	0.97	0.72	0.96	0.70	0.95
All features (12)	0.97	0.75	0.97	0.72	0.96

Table 5
Mean margin per class with respect to type and number of features.

Features	Building	Vegetation	Artificial ground	Natural ground	Mean margin
RGB (3)	0.55	−0.35	0.70	−0.22	0.57
ME lidar (5)	0.86	−0.10	0.87	−0.94	0.78
FW lidar (4)	0.81	0.20	0.78	−0.01	0.75
ME + FW lidar (9)	0.89	0.28	0.82	−0.22	0.80
Selected features (5)	0.89	0.33	0.87	0.31	0.84
All features (12)	0.89	0.35	0.85	0.30	0.84

classification of dense urban scenes since natural ground is a minor class.

5.7. Multi-source feature analysis

In order to confirm the relevance of most important features, classification was run using multispectral image, ME lidar and FW lidar features separately, then using all lidar features, and finally combining optical image and lidar features. The accuracies and the margins are compared to those obtained using the five selected features.

5.7.1. Classification accuracy analysis

Table 4 sums up the accuracy results for all classes as well as per class. The results are averaged over ten classification runs. They confirm the importance of the joint use of multispectral images and lidar data for urban scene classification, and the contribution of FW lidar features which are both highlighted in the table.

When observing the extensive accuracy comparison (*cf.* Table 4), many points are highlighted.

1. The *RGB* bands result in the lowest global classification accuracy and fail to correctly classify the vegetation class. This is due to image sensitivity to the illumination angles (building rooftops), and due to the presence of shadows in dense urban scenes.
2. ME lidar features are suited to building and artificial ground classification. However, they do not allow discriminating between both types of grounds since ME lidar data only provide spatial information. Consequently, natural ground points are misclassified as artificial ground class which is the major class. As for vegetation, only an accuracy of 50% is achieved due to the laser properties in these areas (*cf.* Section 5.4).
3. FW lidar features give better results than ME lidar without using Δz. In fact, the four FW lidar features are complementary, as shown in Table 1. The amplitude A discriminates artificial ground by low values. The echo width w is well suited to vegetation. The echo cross-section σ discriminates between building and artificial ground leading to high and low values respectively, and finally α is important for the natural ground class (*cf.* Fig. 9(d)). This confirms the relevance of FW lidar data for classifying urban scenes.
4. Using all nine lidar features enhances the results especially for buildings and vegetation. Both classes are considered as off-ground which is better discriminated using Δz. However, the results are worse for the natural ground class. As shown in Fig. 9(d), the variable importance is dispersed for this class. Using all lidar features disturbs the training model which results in a lower accuracy.
5. The most relevant five features maintain a satisfactory classification accuracy of 95% while reducing the number of features. When compared to all features' results, accuracies mainly differ for minor classes. As shown in Fig. 9(b) and (d), the feature importance is dispersed. Reducing the number of features affects essentially these minor classes.

5.7.2. Margin analysis

Table 5 compares the mean margin value for different input features. One can observe that the highest margin is obtained with the five selected features. The higher the margin, the stronger the ensemble classifier. A high negative margin value indicates that most trees voted for wrong classes. This may be due to the non-suitability of the input features for the true class or to noisy samples. The joint interpretation of Tables 4 and 5 shows that the RF classifier with the best five features has the highest positive mean margin while keeping a good classification accuracy. This confirms that the five selected features are relevant for urban scene classification. The confidence is even improved in comparison with the RF classifier involving all features for all classes except for vegetation. The former classifier maximizes the minimal margins for these classes. However, the latter classifier performs better for vegetation since two of the remaining features are well suited to vegetation (N_e and R_z, *cf.* Fig. 4).

We can notice that using only ME lidar features for the natural ground class leads to a high negative margin and a poor classification accuracy. The margin value leads to a high confidence in misclassification; in this case, the input features are not suited to the class of interest. Indeed, ME lidar features are spatial features which do not discriminate between both types of grounds.

The results show that FW lidar features are required to distinguish between artificial and natural grounds. In addition, they significantly improve the classification of vegetation. The global accuracy is enhanced in comparison with ME lidar features. The minimal margins for vegetation and natural ground are improved since FW lidar features better characterize the physical properties of objects, but the confidence still remains low with values around 0. However, with FW lidar features, the confidence is lower for major classes due to the non-use of Δz which better discriminates between buildings and artificial ground.

When combining ME and FW lidar features, the confidence is enhanced for all classes except natural ground. For the latter class, the ME lidar features are not suitable as explained before. So they introduce a kind of noise for this class. The joint use of both lidar acquisition modes highly improves the classification accuracy of the vegetation class.

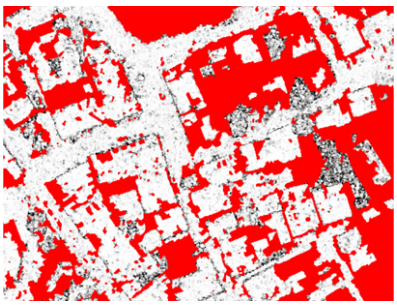
Author's personal copy

(a) RGB orthoimage.

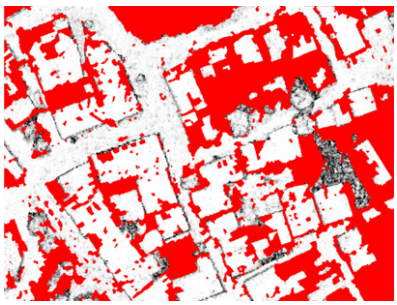
(b) Ground truth.

(c) RGB features, $\bar{m} = 0.57$.

(d) ME lidar features, $\bar{m} = 0.784$.



(e) FW lidar features, $\bar{m} = 0.754$.



(f) ME + FW lidar features, $\bar{m} = 0.805$.

(g) Five selected features, $\bar{m} = 0.844$.

Fig. 10. Margin image comparison depending on multi-source features.

Fig. 10 depicts the corresponding margin images with *RGB*, ME lidar, FW lidar, joint use of ME and FW lidar, and the best five features. With multispectral *RGB* features (Fig. 10(c)), one can observe that margin values are low on shadowed rooftops and grounds leading to the lowest mean margin $\bar{m}$. When using ME lidar features (Fig. 10(d)), the margin values are higher especially for major classes (building and artificial ground). We can see the natural ground class in black color due to the non-suitability of features to this class. For the FW lidar features' margin (Fig. 10(e)), the image presents globally a "salt and pepper" effect which denotes a lower confidence for major classes and still lower values for minor classes even if they are enhanced (*cf.* Table 5). This is due to noisy initial FW features. However, the global accuracy is improved (*cf.* Table 4). The joint use of ME and FW lidar features enhances the classification accuracy for all classes except for natural ground due to ME lidar properties. The confidence measure is improved for building and vegetation classes (*cf.* Table 5). This can be observed from the corresponding image margin (Fig. 10(f)). Indeed, among FW lidar features, the echo amplitude A is well suited to building, whereas echo cross-section σ is adapted for vegetation.

Finally, the mean margin of the RF classifier with five features (Fig. 10(g)) appears less noisy. Margin values are higher for all classes. Small margin values are mostly located on building facades and shadowed areas of vegetation and natural ground.

6. Conclusion

In this paper, we have studied the relevance of multi-source optical image and lidar data features for urban scene classification. Lidar features are composed of multi-echo and full-waveform features. The Random Forests classifier is used to assess both the feature importance measures, and the classifier strength. The permutation accuracy criteria have been used to study the importance of each feature for the whole urban scene and for each class. Through our experiments, several observations have been made: (1) the most significant feature is topographic, the relative height of a lidar point; (2) the best five feature vector is $[\Delta z, B, R, A, \sigma]$ that is composed of two optical image channels, one topographic ME lidar feature, and two FW lidar features. This result shows the relevance of a joint use of airborne lidar data and optical image features; (3) the contribution of FW lidar features is demonstrated in comparison with ME lidar features. The results are enhanced especially when the spatial information is insufficient (*i.e.* natural ground); (4) some features appear to be specific to a particular class such as the normalized number of echoes N_e for the vegetation.

Moreover, the margin theory has been introduced as a measure of confidence and to assess the strength of a RF classifier. We observed that: (1) confidence for the RF classifier with *RGB* features is very low especially in shadowed areas; (2) ME lidar provides a high confidence for building and artificial ground and a very poor

Author's personal copy

confidence for natural ground since the related features are only topographic or geometric; (3) FW lidar enhances the confidence for building and vegetation classes. The RF classifier with the five selected features has the highest mean margin while achieving a good accuracy. This confirms that these features are the most discriminating for urban scene classification. A global classification accuracy of 95% is achieved. However, in dense urban scenes, the data are highly imbalanced, and, therefore the minor classes (vegetation and natural ground) are harder to classify leading to accuracies of respectively 72% and 70%. This result could be improved by a two-pass Random Forests classifier more adapted to minor classes and involving low margins. In addition, using an adaptive 2D neighborhood for lidar features, or radiometric equalization of multispectral image should give higher importance to the corresponding features. This study also showed that some features are not of significance to extract the four urban classes such as planarity related features (N_z, $\mathcal{R}_z$). They should be more useful for roof building segmentation.

The proposed feature importance framework can be applied to select the appropriate features for classification, segmentation or multi-source fusion applications. It can also be applied to remote sensing data such as hyperspectral data which involve hundreds of features, as well as multi-source, multi-sensor data. Finally, this work can also benefit other application fields such as biomedical or computer vision.

References

Breiman, L., 2001. Random forests. Machine Learning 45 (1), 5–32.

Brenner, C., 2010. In: Vosselman, G., Maas, H.-G. (Eds.), Airborne and Terrestrial Laser Scanning. Whittles Publishing, Dunbeath, Scotland, UK.

Carlberg, M., Gao, P., Chen, G., Zakhor, A., 2009. Classifying urban landscape in aerial LiDAR using 3D shape analysis. In: Proceedings of the IEEE International Conference on Image Processing. IEEE, Cairo, Egypt, pp. 1701–1704.

Charaniya, A., Manduchi, R., Lodha, S., 2004. Supervised parametric classification of aerial LiDAR data. In: Proceedings Real-Time 3D Sensors and their Use Workshop, in Conjunction with IEEE CVPR. Washington, DC, USA. 27 June–2 July (on CD-ROM).

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M., Puech, W., 2007. Processing full-waveform LiDAR data: modelling raw signals. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), 102–107.

Chehata, N., Guo, L., Mallet, C., 2009a. Airborne LiDAR feature selection for urban classification using random forests. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 39 (Part 3/W8), 207–212.

Chehata, N., Guo, L., Mallet, C., 2009b. Contribution of airborne full-waveform LiDAR and image data for urban scene classification. In: Proceedings IEEE International Conference on Image Processing (ICIP). IEEE, Cairo, Egypt, pp. 1669–1672.

Gislason, P., Benediktsson, J., Sveinsson, J., 2006. Random forests for land cover classification. Pattern Recognition Letters 27 (4), 294–300.

Gross, H., Jutzi, B., Thoennessen, U., 2007. Segmentation of tree regions using data of a full-waveform laser. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part (3/W49A)), 57–62.

Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2–3), 130–137.

Ham, J., Chen, Y., Crawford, M., Ghosh, J., 2005. Investigation of the random forest framework for classification of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 43 (3), 492–501.

Höfle, B., Hollaus, M., 2010. Urban vegetation detection using high density full-waveform airborne LiDAR data—combination of object-based image and point cloud analysis. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 7B), 281–286.

Mallet, C., Bretar, F., 2009. Full-waveform topographic LiDAR: state-of-the-art. ISPRS Journal of Photogrammetry and Remote Sensing 64 (1), 1–16.

Mallet, C., Bretar, F., Soergel, U., 2008. Analysis of full-waveform LiDAR data for classification of urban areas. Photogrammetrie Fernerkundung GeoInformation (PFG) 2008 (5), 337–349.

Matei, B., Sawhney, H., Samarasekera, S., Kim, J., Kumar, R., 2008. Building segmentation for densely built urban regions using aerial LiDAR data. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Anchorage, AK, USA, pp. 1–8.

Matikainen, L., Hyyppä, J., Hyyppä, H., 2003. Automatic detection of buildings from laser scanner data for map updating. International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences 33 (Part 3/W13), 218–224.

Melzer, T., 2007. Non-parametric segmentation of ALS point clouds using mean shift. Journal of Applied Geodesy 1 (3), 159–170.

Pal, M., 2005. Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26 (1), 217–222.

Poullis, C., Yu, S., 2009. Automatic reconstruction of cities from remote sensor data. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL, USA, pp. 2775–2782.

Reitberger, J., Schnorr, C., Krzystek, P., Stilla, U., 2009. 3ṣegmentation of single trees exploiting full waveform LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing 64 (6), 561–574.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2007. Building detection by fusion of airborne laser scanner data and multi-spectral images: performance evaluation and sensitivity analysis. ISPRS Journal of Photogrammetry and Remote Sensing 62 (2), 135–149.

Rutzinger, M., Höfle, B., Hollaus, M., Pfeifer, N., 2008. Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. Sensors 8 (8), 4505–4528.

Schapire, R., Freund, Y., Bartlett, P., Lee, W., 1998. Boosting the margin: a new explanation for the effectiveness of voting methods. The Annals of Statistics 26 (5), 1651–1686.

Secord, J., Zakhor, A., 2007. Tree Detection in urban regions using aerial LiDAR and image data. IEEE Geoscience and Remote Sensing Letters 4 (2), 196–200.

Sithole, G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing 59 (1–2), 85–101.

Tang, E.K., Suganthan, P.N., Yao, X., 2006. An analysis of diversity measures. Machine Learning 65 (1), 247–271.

Wagner, W., Hollaus, M., Briese, C., Ducic, V., 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. International Journal of Remote Sensing 29 (5), 1433–1452.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing 60 (2), 100–112.

Waske, B., Benediktsson, J., 2007. Fusion of support vector machines for classification of multisensor data. IEEE Transactions on Geoscience and Remote Sensing 45 (12), 3858–3866.

Waske, B., Braun, M., 2009. Classifier ensembles for land cover mapping using multitemporal sar imagery. ISPRS Journal of Photogrammetry and Remote Sensing 64 (5), 450–457.

Xu, G., Zhang, Z., 1996. Epipolar Geometry in Stereo, Motion and Object Recognition. Kluwer Academic Publishers, Boston, MA, USA.

Zhou, Q.-Y., Neumann, U., 2009. A streaming framework for seamless building reconstruction from large-scale aerial LiDAR data. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL, USA, pp. 2759–2766.

Zhu, M., 2008. Kernels and ensembles: perspectives on statistical learning. The American Statistician 62 (2), 97–109.

5.1.3 Sélection de bandes utiles pour la caractérisation de matériaux urbains

EXTRACTION OF OPTIMAL SPECTRAL BANDS USING HIERARCHICAL BAND MERGING OUT OF HYPERSPECTRAL DATA

A. Le Bris[a]*, N. Chehata[b,c], X. Briottet[d], N. Paparoditis[a]

[a] Université Paris-Est, IGN/SR, MATIS, 73 avenue de Paris, 94160 Saint Mandé , France - (arnaud.le-bris, nicolas.paparoditis)@ign.fr
[b] IRD/UMR LISAH El Menzah 4, Tunis, Tunisia
[c] Bordeaux INP, G&E, EA 4592, F-33600, Pessac, France - nesrine.chehata@ensegid.fr
[d] ONERA, The French Aerospace Lab, 2 avenue Edouard Belin, BP 74025, 31055 Toulouse Cedex 4, France - xavier.briottet@onera.fr

KEY WORDS: Classification, Feature selection, Sensor design, Hyperspectral, Spectral bands, Multispectral

ABSTRACT:

Spectral optimization consists in identifying the most relevant band subset for a specific application. It is a way to reduce hyperspectral data huge dimensionality and can be applied to design specific superspectral sensors dedicated to specific land cover applications. Spectral optimization includes both band selection and band extraction. On the one hand, band selection aims at selecting an optimal band subset (according to a relevance criterion) among the bands of a hyperspectral data set, using automatic feature selection algorithms. On the other hand, band extraction defines the most relevant spectral bands optimizing both their position along the spectrum and their width. The approach presented in this paper first builds a hierarchy of groups of adjacent bands, according to a relevance criterion to decide which adjacent bands must be merged. Then, band selection is performed at the different levels of this hierarchy. Two approaches were proposed to achieve this task : a greedy one and a new adaptation of an incremental feature selection algorithm to this hierarchy of merged bands.

1. INTRODUCTION

High dimensional remote sensing imagery, such as hyperspectral imagery, generates huge data volumes, consisting of hundreds of contiguous spectral bands. Nevertheless, most of these spectral bands are highly correlated to each other. Thus using all of them is not necessary. Besides, some difficulties are caused by this high dimensionality, as for instance the curse of dimensionality or data storage problems. To answer these general problems, dimensionality reduction strategies aim thus at reducing data volume minimizing the loss of useful information and especially of class separability. These approaches belong either to **feature extraction** or **feature selection** categories.
Feature extraction methods consist in reformulating and summing up original information, reprojecting it in another feature space. Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discriminant Analysis (LDA) are state-of-the-art feature extraction techniques.
On the opposite, feature selection (FS) methods applied to **band selection** select the most relevant band subset (among the original bands of the hyperspectral data set) for a specific problem.
Furthermore, in the case of hyperspectral data, adjacent bands are very correlated to each other. Thus **band extraction**, that is to say the definition of an optimal set of spectral bands optimizing both their width and position along the spectra, can be considered as intermediate between feature extraction techniques and individual band selection. Band selection/extraction approaches offer advantages compared to feature extraction techniques. First, they make it possible not to loose the physical meaning of the selected bands. Most important, they are adapted to the design of multispectral or superspectral sensors dedicated to a specific application, that is to say sensors designed to deal with specific land cover classification problems for which only a limited band subset is relevant.

*Corresponding author

1.1 Feature selection

Feature selection (FS) can be seen as a classic optimization problem involving both a metric (that is to say a FS score measuring the relevance of feature subsets) to optimize and an optimization strategy.
Even though hybrid approaches involving several criteria exist (Estévez et al., 2009, Li et al., 2011), FS methods and criteria are often differentiated between “filter”, “wrapper” and “embedded”. It is also possible to distinguish supervised and unsupervised ones, whether classes are taken into account.

Filters Filter methods compute a score of relevance for each feature independently from any classifier. Some filter methods are ranking approaches : features are ranked according to a score of importance, as the ReliefF score (Kira and Rendell, 1992) or a score calculated from PCA decomposition (Chang et al., 1999). Other filters associate a score to feature subsets. In supervised cases, separability measures such as Bhattacharyya or Jeffries-Matusita (JM) distances can be used in order to identify the feature subsets making it possible to best separate classes (Bruzzone and Serpico, 2000, Serpico and Moser, 2007). High order statistics from information theory such as divergence, entropy and mutual information can also be used to select the best feature subsets achieving the minimum redundancy and the maximum relevance, either in unsupervised or supervised situations: (Martínez-Usó et al., 2007) first cluster “correlated” features and then select the most representative feature of each group, while (Battiti, 1994, Estévez et al., 2009) select the set of bands that are the most correlated to the ground truth and the less correlated to each other.

Wrappers For wrappers, the relevance score associated to a feature subset corresponds to the classification performance (measured by a classification quality rate) reached using this feature subset. Examples of such approaches can be found in (Estévez et al., 2009, Li et al., 2011) using SVM classifier, (Zhang et al., 2007) using maximum likelihood classifier, (Díaz-Uriarte and

De Andres, 2006) using random forests or even (Minet et al., 2010) for target detection.

Embedded Embedded FS methods are also related to a classifier, but feature selection is performed using a feature relevance score different from a classification performance rate. Some embedded approaches are regularization models associating a fit-to-data term (e.g. a classification error rate) associated to a regularization function, penalizing models when the number of features increases (Tuia et al., 2014). Other embedded approaches progressively eliminate features from the model, as SVM-RFE (Guyon et al., 2002) that considers the importance of the features in a SVM model. Other approaches have a built-in mechanism for feature selection, as Random Forests (Breiman, 2001) that uses only the most discriminative feature among a feature subset randomly selected, when splitting a tree node.

Another issue for band selection is the optimization strategy to determine the best feature subset corresponding to a criteria. An exhaustive search is often impossible, especially for wrappers. Therefore, heuristics have been proposed to find a near optimal solution without visiting the entire solution space. These optimization methods can be divided into incremental and stochastic ones.
Several incremental search strategies have been detailed in (Pudil et al., 1994), including the Sequential Forward Search (SFS) starting from one feature and incrementally adding another feature making it possible to obtain the best score or on the opposite the Sequential Backward Search (SBS) starting for all possible features and incrementally removing the worst feature. Variant such as Sequential Forward Floating Search (SFFS) or Sequential Backward Search (SBFS) are proposed in (Pudil et al., 1994). (Serpico and Bruzzone, 2001) proposes variants of these methods called Steepest Ascent (SA) algorithms.
Among stochastic optimization strategies used for feature selection, several algorithms have been used for feature selection, including Genetic algorithms (Li et al., 2011, Estévez et al., 2009, Minet et al., 2010), Particle Swarm Optimization (PSO) (Yang et al., 2012) or simulated annealing (De Backer et al., 2005, Chang et al., 2011).

1.2 Band grouping and band extraction

Band grouping and clustering In the specific case of hyperspectral data, adjacent bands are often very correlated to each other. Thus, band selection encounters the question of the clustering of the spectral bands of a hyperspectral data set. This can be a way to limit the band selection solution space. Band clustering/grouping has sometimes been performed in association with individual band selection. For instance, (Li et al., 2011) who first group adjacent bands according to conditional mutual information, and then perform band selection with the constraint that only one band can be selected per cluster. (Su et al., 2011) perform band clustering applying k-means to band correlation matrix and then iteratively remove the too inhomogeneous clusters and the bands too different from the representative of the cluster to which they belong. (Martínez-Usó et al., 2007) first cluster "correlated" features and then select the most representative feature of each group, according to mutual information. (Chang et al., 2011) performs band clustering using a more global criterion taking specifically into account the existence of several classes : simulated annealing is used to maximise a cost function defined as the sum, over all clusters and over all classes, of the sum of correlation coefficients between bands belonging to a same cluster. (Bigdeli et al., 2013, Prasad and Bruce, 2008) perform band clustering, but not for band extraction : a multiple SVM classifier is defined, training one SVM classifier per cluster. (Bigdeli et al., 2013) have compared several band clustering/grouping methods, including k-means applied to the correlation matrix or an approach considering the local minima of mutual information between adjacent bands as cluster borders. (Prasad and Bruce, 2008) propose another band grouping strategy, starting from the first band of the spectrum and progressively growing it with adjacent bands until a stopping condition based on mutual information is reached.

Band extraction Specific band grouping approaches have been proposed for spectral optimization. (De Backer et al., 2005) define spectral bands by Gaussian windows along the spectrum and propose a band extraction optimizing score based on a separability criterion (Bhattacharyya error bound) thanks to a simulated annealing. (Cariou et al., 2011) merge bands according to a criteria based on mutual information. (Jensen and Solberg, 2007) merge adjacent bands decomposing some reference spectra of several classes into piece-wise constant functions. (Wiersma and Landgrebe, 1980) define optimal band subsets using an analytical model considering spectra reconstruction errors. (Serpico and Moser, 2007) propose an adaptation of his Steepest Ascent algorithm to band extraction, also optimizing a JM separability measure. (Minet et al., 2010) apply genetic algorithms to define the most appropriate spectral bands for target detection. Last, some studies have also studied the impact of spectral resolution (Adeline et al., 2014), without selecting an optimal band subset.

1.3 Proposed approach

The approach proposed in this paper consists in first building a hierarchy of groups of adjacent bands. Then, band selection is performed at the different levels of this hierarchy. Two approaches are proposed to achieve this task.
Thus, it is here intended to use the hierarchy of groups of adjacent bands as a constraint for band extraction and a way to limit the number of possible combinations, contrary to some existing approaches such as (Serpico and Moser, 2007) that extract optimal bands according to JM information using an adapted optimization method or (Minet et al., 2010) that directly use a genetic algorithm to optimize a wrapper score.

2. DATA SET

The proposed algorithms were mostly tested on the ROSIS VNIR reflectance hyperspectral Pavia Center data set [1]. Its spectral domain ranges from 430nm to 860nm. Its associated land cover ground truth includes the next classes : "water", "trees", "meadows", "self blocking bricks", "bare soil", "asphalt", "roofing bitumen" , "roofing tiles" and "shadows".
They were also tested on the VNIR-SWIR AVIRIS Indian Pines and Salinas scenes, captured over rural areas.

3. HIERARCHICAL BAND MERGING

The first step of the proposed approach consists in building a hierarchy of groups of adjacent bands, that are then merged. Even though it is intended to be used to select an optimal band subset, this hierarchy of merged bands can also be a way to explore several band configuration with varying spectral resolution, that is to say with contiguous bands with different bandwidth.

[1] http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

3.1 Hierarchical band merging algorithm

Notations Let $B = \{\lambda_i\}_{0 \leq i \leq nbands}$ be the original (ordered) set of bands. Let $H = \{H^{(i)}\}_{0 \leq i < nlevels}$ be the hierarchy of merged bands. $H^{(i)} = \{H_j^{(i)}\}_{1 \leq j \leq n_i}$ is the ith level of this hierarchy of merged bands. It is composed of n_i merged bands, that is to say n_i ordered groups of adjacent bands from B.
Thus, each $H_j^{(i)}$ is defined as a spectral domain :
$H_j^{(i)} = [H_j^{(i)}.\lambda_{min}; H_j^{(i)}.\lambda_{max}]$.
Thus, the merged band $B_1 \oplus B_2$ obtained when merging two such adjacent merged bands B_1 and B_2 is $B_1 \oplus B_2 = [B_1.\lambda_{min}; B_2.\lambda_{max}]$
Let $J(.)$ be the score that has to be optimized during the band merging process.

The proposed hierarchical band merging approach is a bottom-up one. The algorithm is defined below :

Initialization : $H^{(0)} = B$ (that is to say that merged band of the first level of the hierarchy only contains one individual original band).
Band merging : create level l+1 from level l :
Find the pair of adjacent bands at level l that will optimize the score if they are merged : find $\hat{k} = argmin_k J(T(H^{(l)}, k))$ with $T(H^{(l)}, k) = [H_0^{(l)}; ...; H_{k-1}^{(l)}; H_k^{(l)} \oplus H_{k+1}^{(l)}; H_{k+2}^{(l)}; ...; H_{n_l}^{(l)}])$
Then $H^{(l+1)} = T(H^{(l)}, \hat{k})$

A table L_l^{l+1} is defined to link the different merged bands at consecutive hierarchy levels :
for $1 \leq j \leq \hat{k}$, $L_l^{l+1}(H_j^{(l)}) = H_j^{(l+1)}$
$L_l^{l+1}(H_{\hat{k}}^{(l)}) = H_{\hat{k}}^{(l+1)}$
$L_l^{l+1}(H_{\hat{k}}^{(l)} + 1) = H_{\hat{k}}^{(l+1)}$
for $\hat{k} + 2 \leq j \leq n_l$, $L_l^{l+1}(H_j^{(l)}) = H_{j-1}^{(l+1)}$

3.2 Band merging criteria

Several optimization scores J can be examined. (In the algorithm described in section 3.1, this score is aimed to be minimized.) They can be either supervised or unsupervised, depending whether classes are considered or not at this step.

3.2.1 Correlation between bands

Between band correlation (either the classic normalized correlation coefficient or mutual information) (see figure 1) measures the dependence between bands. So a first band merging criterion intends to merge adjacent bands considering how they are correlated to each other. Thus, it tries to obtain consistent groups of adjacent correlated bands.
Such measure inspired from (Chang et al., 2011) can be defined by next function (intended to be minimized):
$J(H^{(l)}) = \sum_{i=1}^{n_l} \sum_{b_1=H_i^{(l)}.\lambda_{min}}^{H_i^{(l)}.\lambda_{max}} \sum_{b_2=H_i^{(l)}.\lambda_{min}}^{H_i^{(l)}.\lambda_{max}} (1-c(b1, b2))$
with $c(b_1, b_2)$ the correlation score between bands b_1 and b_2.

3.2.2 Spectra approximation error

Band merging could also use (Jensen and Solberg, 2007)'s method to decompose some reference spectra of several classes into piece-wise constant functions (fig. 2). Adjacent bands are then merged trying to minimize the reconstruction error between the original and the piece-wise constant reconstructed spectra.
Such measure is defined by next function for a set $s_{j\,1 \leq j \leq n_s}$ of n_s spectra :
$J(H^{(l)}) = \sum_{j=1}^{n_s} \sum_{i=1}^{n_l} \sum_{b=H_i^{(l)}.\lambda_{min}}^{H_i^{(l)}.\lambda_{max}} |s_j(b) - mean(s_j, H_i^{(l)})|$
where $mean(s_j, H_i^{(l)})$ denotes the mean of spectra s_j over spectral domain $H_i^{(l)}$

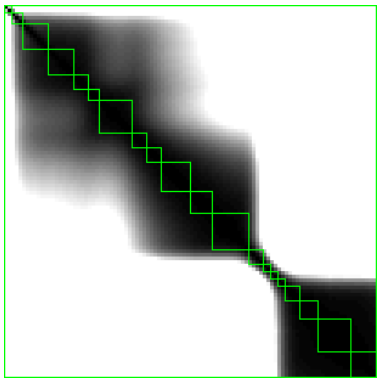


Figure 1: Examples of groups of bands superimposed on the between band correlation matrix (for Pavia data set)

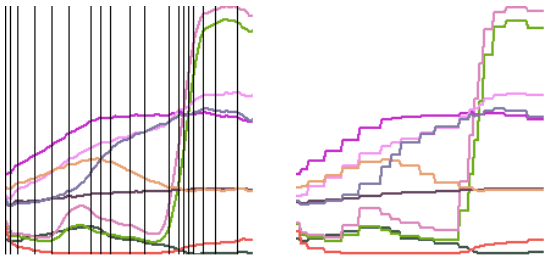


Figure 2: On the left, examples of merged bands superimposed on the original reference spectra. On the right, piece-wise constant reconstructed spectra for these merged bands (Pavia data set)

3.2.3 Separability

Another criterion to merge adjacent band is their contribution to separability between classes. Possible separability measures are the Bhattacharyya distance (*B-distance*) or the Jeffries-Matusita distance (Bruzzone and Serpico, 2000, Serpico and Moser, 2007).
The Bhattacharyya separability between classes i and j is defined as
$B_{ij} = \frac{1}{8}\,{}^t(\vec{\mu}_i - \vec{\mu}_j)\Sigma^{-1}(\vec{\mu}_i - \vec{\mu}_j) + 0.5 \ln\left(\frac{\det \Sigma}{\sqrt{\det \Sigma_i \ \det \Sigma_j}}\right)$,
where $\Sigma = \frac{\Sigma_i + \Sigma_j}{2}$, with $\vec{\mu}_i$ and Σ_i be the mean vector and covariance matrix of class i radiometric distribution. As Bhattacharyya separability is defined for binary problems, its mean over all possible pairs of classes can be used as a global separability measure.
Jeffries-Matusita measure for c classes is then defined as $JM = \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} (1 - e^{-B_{i;j}})$.

At a level of the band merging hierarchy, the best set of merged bands is the one that maximizes class separability. So a possible criterion J (to minimize) for band merging can be defined as $J(H^{(l)}) = -JM(H^{(l)})$

3.3 Results

Obtained results on Pavia data set for the 3 criteria described in previous section can be seen on figure 3. The separability based criterion tends to lead to more different results than the other ones. It can be seen that the different criteria don't consider the same parts of the spectrum as having to be kept at fine resolution. For instance, correlation or spectra reconstruction criteria tend to fast merge bands between number 30 and 32, while separability tend to preserve them at fine resolution. On the opposite, separability tends to fast merge some bands in the red-edge domain, while the other criteria keep this domain at fine resolution. This

can be understood considering the underlying criteria ; indeed adjacent bands are not very correlated to each other in this domain and the slope of spectra is strong for vegetation classes, and thus they not be merged easily according to correlation or spectra approximation error band merging criteria. On the opposite, the only interesting information for classification (e.g. for class separability) is the fact there is a slope there and thus the values of the bands before and after this domain. Thus, merging these red-edge bands will have little impact on class separability.

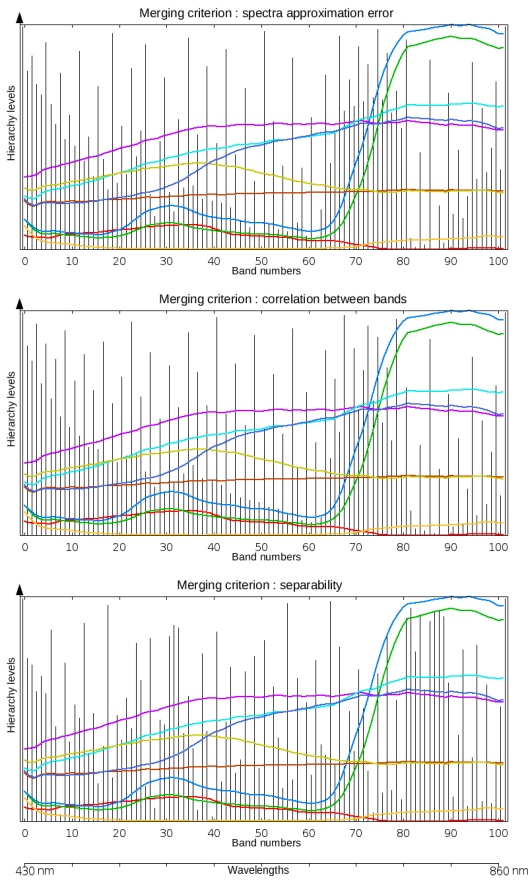


Figure 3: Hierarchies of merged bands obtained for different criteria for Pavia data set: spectra piece-wise approximation error (top), between band correlation (middle) and class separability (bottom). x-axis corresponds to the band numbers/wavelengths. y-axis corresponds to the level in the band merging hierarchy (bottom : finest level with original bands, top : only a single merged band). Vertical black lines are the limits between merged bands : the lower in the hierarchy, the more merged bands. Reference spectra of the classes are displayed in colour.

As the hierarchy of merged bands can also be a way to explore several band configuration with varying contiguous bands with different spectral resolution, the different band configurations corresponding to the different levels were evaluated using a classification quality measure. Thus, for each level, a classification was performed using a support vector machine (SVM) classifier with a radial basis function (rbf) kernel and evaluated. Its Kappa coefficient was considered.
Such results are presented on figure 4. It can be seen that some spectral configurations made it possible to obtain better results than at original spectral resolution. Configurations obtained using the correlation coefficient are generally less good than for the two other criteria. Except for Pavia, the spectra piece-wise approximation error merging criterion tends to lead to the best results. But for Pavia, the classification Kappa reached using the different criteria remained very similar.

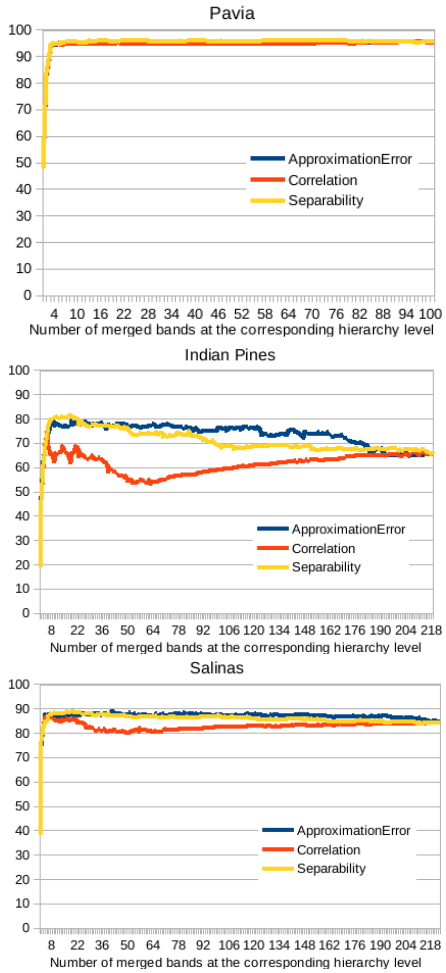


Figure 4: Kappa (in %) reached by a rbf SVM for the different band configurations of the hierarchy (x-axis = number of merged bands in the spectral configuration corresponding to the hierarchy level), for Pavia (top), Indian Pines (middle) and Salinas (bottom) data sets.

4. BAND SELECTION USING A GREEDY METHOD

To optimize spectral configuration for a limited number of merged bands, a greedy approach was first used : it performed band selection at the different levels of the hierarchy of merged bands, paying no attention at results obtained at the previous level. Thus a set of merged bands was selected at each level of the hierarchy. The feature selection (FS) score to optimize was the Jeffries-Matusita separability measure. It was optimized at each level of the hierarchy using an incremental optimization heuristic called Sequential Forward Floating Search (SFFS) (Pudil et al., 1994) and reminded below in its general formulation.

4.1 Sequential Forward Floating Search

It is intended to select less than p features among a feature set B. Let S be the selected band subset and J the FS score to maximize.

Initialization : Find band b in B such that $b = argmax_{z \in B} J(\{z\})$
$S \leftarrow \{b\}$
$n \leftarrow 1$

$J_1 \leftarrow J(S)$
While $\#S < p$
Find band $b \in B \setminus S$ such that $S \cup \{b\}$ maximizes the FS score, i.e. $b = argmax_{z \in B \setminus S} J(S \cup \{z\})$
$S \leftarrow S \cup \{b\}$
Question S : find band $s \in S$ such that $S \setminus \{s\}$ maximizes FS score, i.e. $s = argmax_{z \in S} J(S \setminus \{z\})$. This means that s is less important than the other bands of S, since removing it decreases the FS score less.
if $s = b$
$J_n \leftarrow J(S)$
$n \leftarrow n + 1$
else
$S \leftarrow S \setminus \{s\}$
while$J(S) > J_{n-1}$
$n \leftarrow n - 1$
$J_n \leftarrow J(S)$
Question S : find band $s \in S$ such that $S \setminus \{s\}$ maximizes FS score, i.e. $s = argmax_{z \in S} J(S \setminus \{z\})$.
endwhile
endif
endwhile

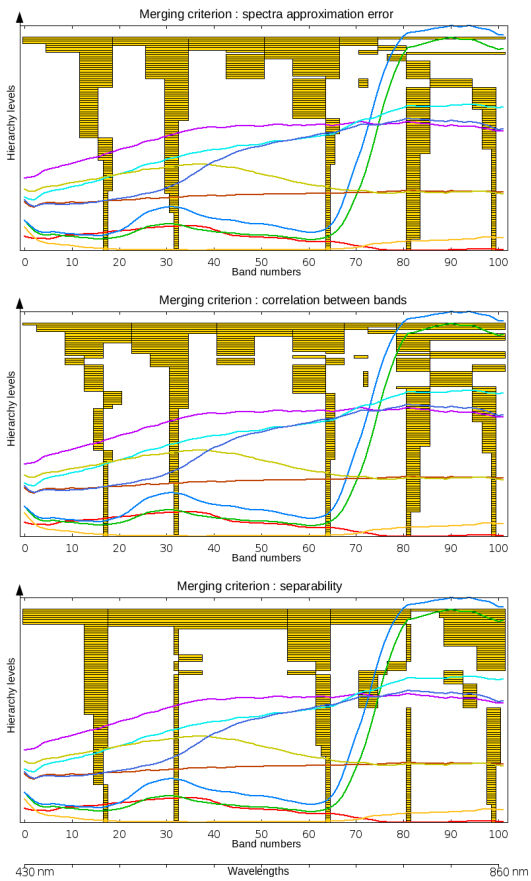


Figure 5: Pavia data set: selected bands at the different levels of the hierarchy using the greedy approach for hierarchies of merged bands obtained using different band merging criteria : spectra piece-wise approximation error (top), between band correlation (middle), class separability (bottom). x-axis corresponds to the band numbers/wavelengths. y-axis corresponds to the level in the band merging hierarchy (bottom : finest level with original bands, top : only a single merged band).

4.2 Results

Obtained results on Pavia data set are presented on figure 5 : 5 merged bands (as in (Le Bris et al., 2014)) were selected at each level of the hierarchy of merged bands. It can be seen that the positions of the selected merged bands don't change a lot when climbing the hierarchy, except when reaching the lowest spectral resolution configurations. It can also be noticed that at some level of the hierarchy the position of some selected merged bands can move and then come back to its initial position when climbing the hierarchy.
Thus, it can be possible to use the selected bands at a level l to initialize the algorithm at next level $l + 1$. This modified method will be presented in section 5..

The merged band subsets selected at the different levels of the hierarchy were evaluated according to a classification quality measure. As in previous section, the Kappa coefficient reached by a rbf SVM was considered. Results for Pavia and Indian Pines data sets can be seen on figure 6. At each level of the hierarchy, 5 bands were selected for Pavia, and 10 bands for Indian Pines. It can be seen that these accuracies remain very close to each other whatever the band merging criterion used, and no band merging criterion tend to really be better than the other ones. Results obtained using merged bands are generally better than using the original bands.

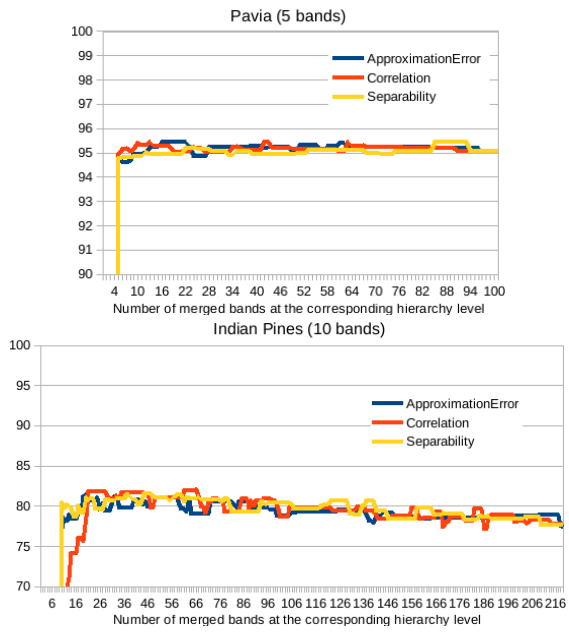


Figure 6: Kappa (in %) reached for rbf SVM classification for merged band subsets selected at the different levels of the hierarchy for Pavia and Indian Pines data sets using the greedy FS algorithm (x-axis = number of merged bands in the spectral configuration corresponding to the hierarchy level).

5. TAKING INTO ACCOUNT THE BAND MERGING HIERARCHY DURING FEATURE SELECTION

5.1 Algorithm

Previous merged band selection approach is greedy and computing time expensive. So an adaptation of the SFFS heuristic was proposed to directly take into account the band merging hierarchy in the band selection process. As for the hierarchical band merging algorithm, a bottom-up approach was chosen. Contrary

	Original bands	Greedy SFFS	Adapted SFFS
Pavia (5 bands)			
Kappa (%)	95.05	95.45	95.44
Computing times	2min	1h10min	9min
Indian Pines (10 bands)			
Kappa (%)	77.69	81.41	81.21
Computing times	4min	7h	40min

Figure 7: Computing times and best Kappa coefficients reached on Pavia (for a 5 band subset) and Indian Pines (for a 10 band subset) data sets for band merging criterion "spectra piece-wise approximation error"

to the greedy approach, this new algorithm uses the band subset selected at the previous lower level when performing band selection at a new level of the hierarchy of merged bands.
This algorithm is described below :
Let $S^{(l)} = S_i^{(l)}{}_{1 \leq i \leq p}$ be the set of selected merged bands at level l of the hierarchy. (NB : The same number p of bands is selected at each level of the hierarchy.)
Initialization : standard SFFS band selection algorithm is applied to the base level $H^{(0)}$ of the hierarchy
Iterations over the levels of the hierarchy :
Generate $S^{(l+1)}$ from $S^{(l)}$:
$S^{(l+1)} \leftarrow L_l^{l+1}(S_i^{(l)})_{1 \leq i \leq p}$
Remove possible duplications from $S^{(l+1)}$
if $\#S^{(l+1)} < p$,
find $s = argmax_{b \in H^{(l+1)} \setminus S^{(l+1)}} J(S^{(l+1)} \cup b$
$S^{(l+1)} \leftarrow S^{(l+1)}; s$
endif
Question $S^{(l+1)}$: find band $s \in S^{(l+1)}$ such that $S^{(l+1)} \setminus \{s\}$ maximizes FS score, i.e. $s = argmax_{z \in S^{(l+1)}} J(S^{(l+1)} \setminus \{s\})$.
$S^{(l+1)} \leftarrow S^{(l+1)} \setminus s$
Then apply classic SFFS algorithm until $\#S^{(l+1)} = p$.

5.2 Results

Obtained results on Pavia scene for the band merging criterion "spectra piece-wise approximation error" are presented on figure 8 : 5 merged bands were selected at each level of the hierarchy, starting from an initial solution obtained at the bottom level of the hierarchy.
As for previous experiments, obtained results were evaluated both for Pavia (5 selected bands) and Indian Pines (10 selected bands) data sets. Kappa reached for rbf SVM classification for merged band subsets selected at the different levels of the hierarchy (built for band merging criterion "spectra piece-wise approximation error") can be seen both for the greedy FS algorithm and for the hierarchy aware one on figure 9 : obtained results remain very close, whatever the optimization algorithm.
It can be said from table 7 that both algorithms lead to equivalent results considering classification performance while the proposed hierarchy aware algorithm is really faster.

6. CONCLUSION

In this paper, a method was proposed to extract optimal spectral band subsets out of hyperspectral data sets. A hierarchy of merged bands was first built according to a band merging criterion. It was then used to explore the solution space for band extraction : band selection was then performed at each level of the hierarchy, either using a greedy approach or an adapted hierarchy aware approach. Classification results tend to be slightly improved when using merged bands, compared to a direct use of the original bands. Besides, in the context of band optimization for sensor design, it can also be a way to get more photons.

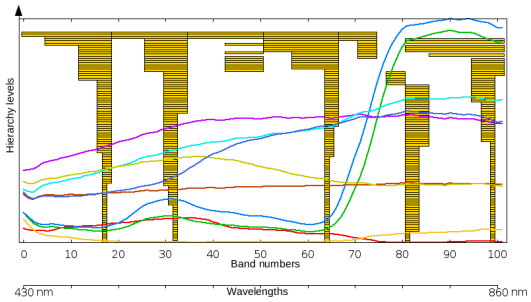


Figure 8: Pavia data set: selected bands at the different levels of the hierarchy using the proposed hierarchy aware algorithm for a hierarchy of merged bands obtained using spectra piece-wise approximation error band merging criteria

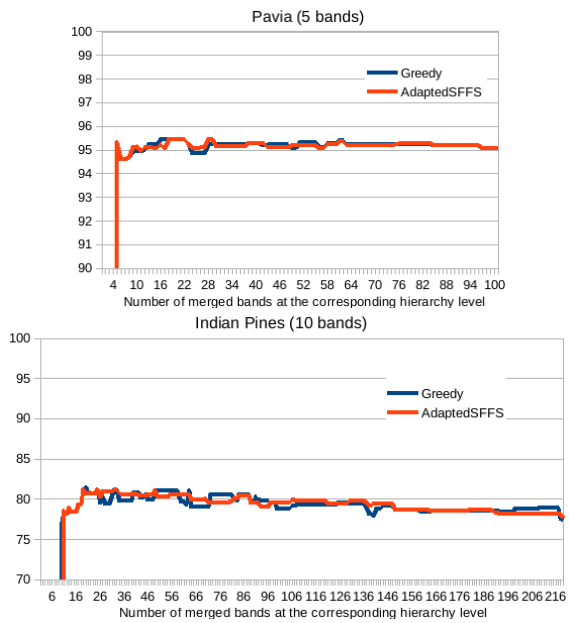


Figure 9: Kappa (in %) reached for rbf SVM classification for merged band subsets selected at the different levels of the hierarchy (built for band merging criterion "spectra piece-wise approximation error") for Pavia and Indian Pines data sets, using the hierarchy aware band selection algorithm.

Further work will investigate band optimization aiming at selecting merged bands at different levels of the hierarchy.
This method will also be applied to a specific sensor design band optimization problem : optimizing spectral bands for urban material classification within the French ANR HYEP ANR 14-CE22-0016-01 project.

REFERENCES

Adeline, K., Gomez, C., Gorretta, N. and Roger, J., 2014. Sensitivity of soil property prediction obtained from vnir/swir data to spectral configurations. In: Proc. of the 4th International Symposium on Recent Advances in Quantitative Remote Sensing: RAQRS'IV.

Battiti, R., 1994. Using mutual information for selecting features in supervised neural net learning. IEEE Transactions on Neural Networks.

Bigdeli, B., Samadzadegan, F. and Reinartz, P., 2013. Band grouping versus band clustering in svm ensemble classification of hyperspectral imagery. Photogrammetric Engineering and Remote Sensing 79(6), pp. 523–533.

Breiman, L., 2001. Random forests. Machine Learning 45(1), pp. 5–32.

Bruzzone, L. and Serpico, S. B., 2000. A technique for feature selection in multiclass problem. International Journal of Remote Sensing 21(3), pp. 549–563.

Cariou, C., Chehdi, K. and Le Moan, S., 2011. Bandclust: an unsupervised band reduction method for hyperspectral remote sensing. IEEE Geoscience and Remote Sensing Letters 8(3), pp. 565–569.

Chang, C.-I., Du, Q., Sun, T.-L. and Althouse, M., 1999. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing 37(6), pp. 2631–2641.

Chang, Y.-L., Chen, K.-S., Huang, B., Chang, W.-Y., Benediktsson, J. and Chang, L., 2011. A parallel simulated annealing approach to band selection for high-dimensional remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4(3), pp. 579–590.

De Backer, S., Kempeneers, P., Debruyn, W. and Scheunders, P., 2005. A band selection technique for spectral classification. IEEE Geoscience and Remote Sensing Letters 2(3), pp. 319–232.

Díaz-Uriarte, R. and De Andres, S. A., 2006. Gene selection and classification of microarray data using random forest. BMC bioinformatics 7(3), pp. 1–13.

Estévez, P. A., Tesmer, M., Perez, C. A. and Zurada, J. M., 2009. Normalized mutual information feature selection. IEEE Transactions on Neural Networks 20(2), pp. 189–201.

Guyon, I., Weston, J., Barnhill, S. and Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Machine Learning 46, pp. 289–422.

Jensen, A.-C. and Solberg, A.-S., 2007. Fast hyperspectral feature reduction using piecewise constant function approximations. IEEE Geoscience and Remote Sensing Letters 4(4), pp. 547–551.

Kira, K. and Rendell, L., 1992. A practical approach to feature selection. In: Proceedings of the 9th International Workshop on Machine Learning, pp. 249–256.

Le Bris, A., Chehata, N., Briottet, X. and Paparoditis, N., 2014. Identify important spectrum bands for classification using importances of wrapper selection applied to hyperspectral. In: Proc. of the 2014 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM'14).

Li, S., Wu, H., Wan, D. and Zhu, J., 2011. An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine. Knowledge-based Systems 24, pp. 40–48.

Martínez-Usó, A., Pla, F., Martínez Sotoca, J. and García-Sevilla, P., 2007. Clustering-based hyperspectral band selection using information measures. IEEE Transactions on Geoscience and Remote Sensing 45(12), pp. 4158–4171.

Minet, J., Taboury, J., Pealat, M., Roux, N., Lonnoy, J. and Ferrec, Y., 2010. Adaptive band selection snapshot multispectral imaging in the vis/nir domain. Proceedings of SPIE the International Society for Optical Engineering 7835, pp. 10.

Prasad, S. and Bruce, L. M., 2008. Decision fusion with confidence-based weight assignment for hyperspectral target recognition. IEEE Transactions on Geoscience and Remote Sensing 46(5), pp. 1448–1456.

Pudil, P., Novovicova, J. and Kittler, J., 1994. Floating search methods in feature selection. Pattern Recognition Letters 15, pp. 1119–1125.

Serpico, S. B. and Bruzzone, L., 2001. A new search algorithm for feature selection in hyperspectral remote sensing images. IEEE Transactions on Geoscience and Remote Sensing 39, pp. 1360–1367.

Serpico, S. B. and Moser, G., 2007. Extraction of spectral channels from hyperspectral images for classification purposes. IEEE Transactions on Geoscience and Remote Sensing 45(2), pp. 484–495.

Su, H., Yang, H., Du, Q. and Sheng, Y., 2011. Semisupervised band clustering for dimensionality reduction of hyperspectral imagery. IEEE Geoscience and Remote Sensing Letters 8(6), pp. 1135–1139.

Tuia, D., Volpi, M., Dalla Mura, M., Rakotomamonjy, A. and Flamary, R., 2014. Automatic feature learning for spatio-spectral image classification with sparse svm. IEEE Transactions on Geoscience and Remote Sensing 52(10), pp. 6062–6074.

Wiersma, D. and Landgrebe, D., 1980. Analytical design of multispectral sensors. IEEE Transactions on Geoscience and Remote Sensing GE-18(2), pp. 180–189.

Yang, H., Du, Q. and Chen, G., 2012. Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(2), pp. 544–554.

Zhang, L., Zhong, Y., Huang, B., Gong, J. and Li, P., 2007. Dimensionality reduction based on clonal selection for hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 45(12), pp. 4172–4186.

5.2 Comparaison de données optiques et LiDAR

Comparison of Pleiades and LiDAR Digital Elevation Models for terraces detection in farmlands

Giulia Sofia*, Jean-Stéphane Bailly†, Nesrine Chehata‡, Paolo Tarolli*, Florent Levavasseur§

*University of Padova, Department Land, Environment, Agriculture and Forestry, I-35020 Legnaro
†AgroParisTech, UMR LISAH, F-34060 Montpellier
‡Bordeaux INP, EA 4592, F-33607 Pessac
§Inra, UMR LISAH, F-34060 Montpellier

***Abstract*—Among the most evident anthropogenic modifications of the lanscape, terraces related to agricultural activities are ubiquitous structures that constitute important investments worldwide, and they recently acquired a new relevance to modern concerns about land use management and erosion control. Conservation agriculture and terraces management are an application with great potentialities for Satellite Earth observation and the derived high resolution topography. Due to its high agility, the Pleiades satellite constellation provides new, high-resolution Digital Elevation Models (DEMs) with a sub-metric resolution that could be potentially useful for this task, and their application in a farmland context is nowadays an open research line. This work provides a first analysis, performing an automatic terrace mapping from DEMs obtained from Pleiades images, as compared to LiDAR DEMs. Two existing methods are considered, the Fast Line Segment Detector (LSD) algorithm, and a geomorphometric method based on surface curvature. Despite the lower performances of Pleiades DEMs respect to the LiDAR models, the results indicate that the Pleiades models can be used to automatically detect terrace slopes greater than 2 m with a detection rate of more than 80% of the total length of the terraces. In addition, the results showed that when using noisy digital elevation models, the geomorphometric method is more robust, and it slightly outperforms the LSD algorithm. These results provide a first analysis on how effective Pleiades DEMs can be as an alternative to LiDAR DEMs, also highlighting the future challenges for monitoring large extents in a farmland context.**

***Index Terms*—Stereo-photogrammetry, cultivated landscapes, LiDAR , line detection, accuracy**

I. Introduction

In the past few decades environmental research has become aware of the extent of the human impact on ecologic and geomorphic systems [1]–[5].

Among the most evident human-induced landscape modifications, terraces related to agricultural activities are ubiquitous features, and they constitute important capital investments in a range of landscapes worldwide [4], typically enduring over many human generations [6]. The importance of terraced landscapes has been underlined from different point of views, ranging from cultural and environmental (e.g. the World Cultural Heritage Sites by UNESCO and [7]), to sustainable land use [6], [8], [9], ecosystem importance [10], [11], erosion control, runoff and water management [4], [12], [13], [19], [21]. In recent times, terraced areas acquired a new relevance to modern concerns about land use management and erosion control [15]–[18], being the agricultural land mostly threatened by abandonment [4], or, on the other hand, intensification and specialization of agriculture resulting in heavy land levelling and/or construction of more landslide-prone bench terraces [14]. Giving the terraces obvious relevance public authorities, land managers, and researchers have called for the development of cost-effective and flexible methods for the identification and monitoring of these features [20], as a basis for a correct monitoring, and to diagnose erosion and hydrological risks at the catchment or hillslope scale [4]. Despite their importance, only few terraces have been delineated and are available in national geographical databases [4], [22]. Thus, most of the recent literature dealing with terraces approached to their identification through field survey, interviews with the local population, or interpretation of aerial photographs and successive digitalization (e.g. [6], [23], [24]). Currently, only few automatic methods have been presented to map agricultural terraces [4], [20], [25], [26].

Pleiades is the most recent satellite mission providing optical images at any point of the Earth surface [27]. Due to the agility of the sensor, the Pleiades constellation allows the acquisition of stereo pairs and triplets for highly coherent conditions [28], [29]. The agility of the satellite, and the availability of this dataset have risen questions about what is feasible, and what is efficient, regarding 3D mapping from multi-angle images sequences [27]. While the effectiveness of high resolution datasets in anthropogenic environments has already been proven (see [30] for a full review), the feasibility of Pleiades datasets in the form of Digital Surface Models (DSMs) has been proven for building extractions e.g. [27], but no work in literature has proven so far the effectiveness of Pleiades DEM for agricultural landscapes. Precision farming and agricultural control are application with great potentialities for Earth observation if efficient methodologies could be used to extract useful and accurate end users information [31], and the use of Pleiades DEMs as an alternative to LiDAR DEMs for monitoring large extents and surface changes has yet to be proven and operationally implemented.

Three questions are still unanswered: i) Are Pleiades derived

DEMs accurate and detailed enough for mapping farmland terraces? ii) Given the same resolution, are the mapping performances comparable with LiDAR DEMs? iii) Is there a significant difference in terrace mapping performances when using a Digital Terrain Model (DTM) or a Digital Surface Model (DSM)?

The objective of this paper is to assess the performances of automatic terrace mapping from DEMs obtained from a stereo pair of Pleiades images compare to LiDAR DEMs. The dataset used for the performance assessment covers a 4 km^2 region of gentle, hilly Mediterranean vegetated farmland in southern France. First, the paper presents the datasets and some preliminary analysis results from the DEM comparison. The DEM processing methods, including DEM smoothing, feature extraction and vector post-processing, are discussed next. Then, the performances of both methods and the reliability of the Pleiades DEM for delineating terraces are discussed.

II. Datasets

Reference terrace data, Pleiades images and a LIDAR 1 m DSM and DTM were acquired for a portion of the Peyne catchment (southern France) (Fig. 1). The area is mainly covered by vineyards, with small areas of cereal fields and shrubs. The elevation for the study site ranges between 50 m a.s.l. to about 125 m a.s.l.. The climate in this region is sub-humid Mediterranean, with 600 mm y^{-1} of precipitation. Two short rainy seasons, one in the autumn and one in the spring, occur with intense rainfall. The climate and intensive vine cultivation make this area sensitive to flash flooding and erosion. Consequently, settlements from Roman times built several structures to limit floods and soil erosion, including ditches, embankments and terraces (e.g. [32]). The study site has been intensively surveyed for a long time, having been a part of an environmental observatory since 1992 (ORE OMERE: http://www.obs-omere.org/). In the study area, terraces have heights ranging from approximately 0.3 m to 10 m. Their location and shape have not changed during the period of the different data acquisition dates: there were no significant modification in between 2002 (time of the LiDAR collection), and 2013 (Pleiades survey) (Fig. 1).

A. Ground truth data

A systematic field survey was conducted in 2010, with a survey rate of 1.5-3 km^2 d^{-1} per person, depending on the accessibility of the terrain. The terraces were delineated with an elevation accuracy of 0.1 m for heights lower than 3 m and 0.5 m for elevations higher than 3 m. Regarding the terraces height, the considered reference field survey was in some instances limited by the accessibility of the area, thus a higher accuracy cannot be expected. Because of this, in this study ground height data are used to group the field-surveyed terraces in different height classes, to project the detection results regarding the terrace height. The overall survey quality is comparable to the official information generally available about terraces locations. The surveyed features, including their location and height, were further verified and digitalized using a 0.5 m resolution aerial photo (BD-Ortho ©IGN) (Fig. 2).

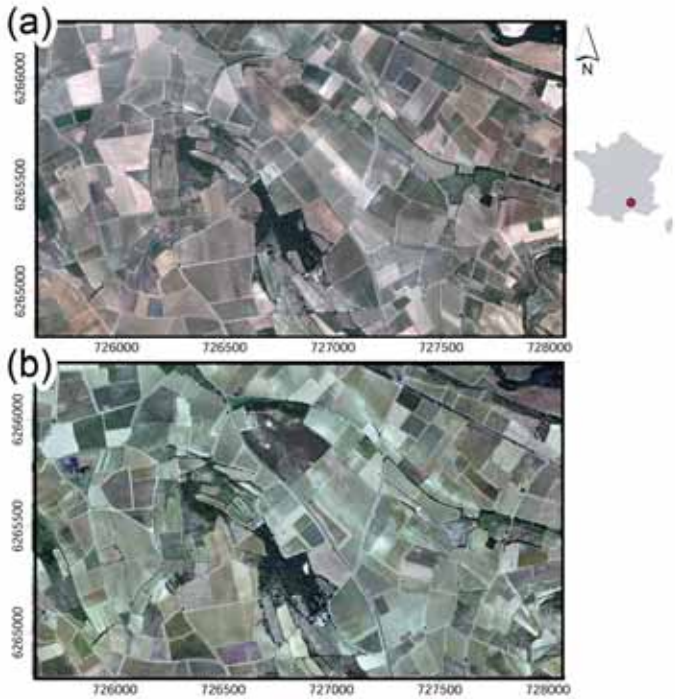


Fig. 1. Location of the study site and view of the area in 2002 (a), and 2013 (b). Images as seen on Google Earth

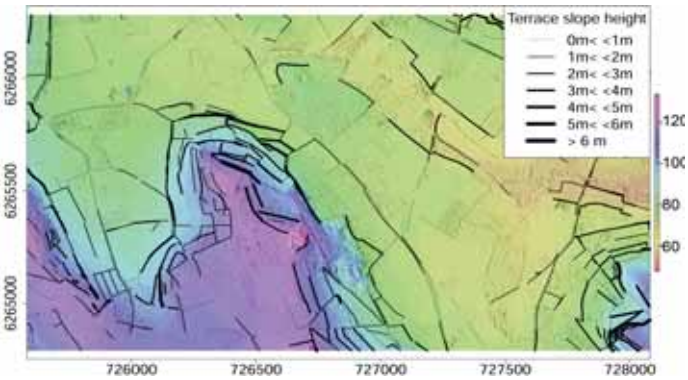


Fig. 2. Ground truth dataset overlapped to the Pleiades DSM

B. Pleiades DEMs

A Digital Surface Model (DSM) was built from a stereo pair of Pleiades images acquired during the leaves-off period of vegetation (January 2013) at a high global incidence angle (30 degrees) leading to a Base to Height (B/H) ratio of $\sim 1/1.6$ which is a usual ratio for stereo pair of satellite images. The DSM was produced by using the open source Mic-Mac (Multi Image Matches for Auto Correlation Methods) tool developed at the French Mapping Agency (IGN) [33], [34].

MicMac is based on the minimization of a global energy that combines a data term which is the image matching score and a regularization term.

$$E_\alpha(Z) = 1 - corr(x, y, Z(x, y)) + \alpha F(\overrightarrow{G}(Z)) \qquad (1)$$

where corr is the normalized cross correlation score at the image projections of point (x, y, Z), $F(\overrightarrow{G}(Z))$ (Eq. 2) is a

positive function that depends on the variations of Z, and it is a regularization term that expresses the a priori knowledge of the surface regularity. α is a weighting parameter for the regularization term.

$$F(\overrightarrow{G}(Z)) = |(Z(x+1,y)-Z(x,y))|+|(Z(x,y+1)-Z(x,y))| \quad (2)$$

The regularization term is well suited for high spatial resolution images since it varies with the terrain discontinuities [35]. In this study, the α parameter was set to be very low to maintain the terrace slopes. To deal with large areas, the MicMac method has a multi-resolution pyramidal approach, that consists in starting the computation at a coarse resolution to reduce the height search space and improving recursively the matching process at each resolution, similar to [36]. The final DSM was resampled at a resolution of 1 m, to be comparable with the resolution of the LiDAR DEMs.

From this DSM, a DTM was generated. Terrace walls present a high local slope, and many DTM filters aim to smooth DSMs (i.e., remove slopes). Consequently, instead of using the usual slope-based filtering algorithm [37], we applied a normal closing filter with circular structural elements of 10 m followed by a Gaussian smoothing filter with a width of 5 m. In mathematical morphology, the closing tends to enlarge the boundaries of foreground (bright) regions in an image (and shrink background color holes in such regions). The effect of the operator is to preserve background regions that have a similar shape to its structuring element, or that can completely contain the structuring element, while eliminating all other regions of background pixels. The choice of the structural element size was governed by the size of the removed object (i.e., vine lines, isolated trees and hedgerows). A Gaussian smoothing filter was then applied to reduce noise on the resulting map while preserving edges on the image. Consequently (Fig. 3), the obtained DTM looks less noisy than the initial DSM, and the terraces are preserved in the landscape.

The final datasets (Pleaides DSM and DTM) show no systematic error over permanent terrain structures (1.8 e-06 m), while presenting an overall standard deviation of errors of 0.51 m with respect to centimetric DGPS points.

C. Multi-echo LiDAR DEMs

LiDAR data were acquired over the study area in June 2002 with a Falcon II Toposys LiDAR system mounted on a helicopter (flight elevation: 900 m; laser pulse emission rate: 83 MHz; 3D points spatial sampling rate: 10 pts m^{-2} [38]). The 3D point cloud allowed the creation of a 1 m DTM and DSM.

The DSM was computed using bi-linear interpolation of the first pulse points at the regular grid DTM node locations. The DTM was computed using the same process from the last pulse scatter of points. From the latter pulse, an additive erosion filter with an adaptive structural element size was applied to remove some residual surface objects (houses, dense forests and hedgerows). The post survey accuracy on the LiDAR last-pulses topographic points on flat and non vegetated areas exhibited a Gaussian noise, with a standard deviation of approximately 0.06 m [38] with respect to centimetric DGPS points.

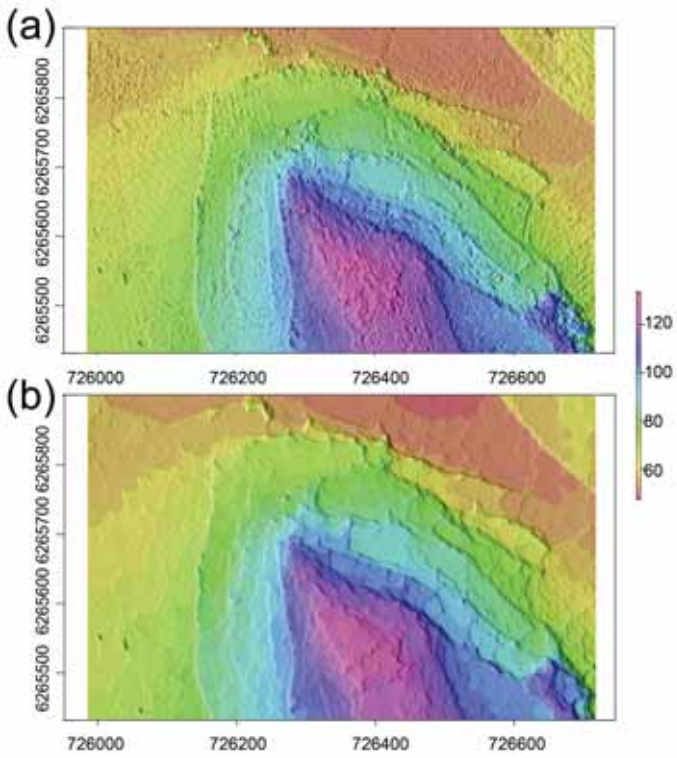


Fig. 3. Pleiades DSM to DTM process: initial DSM (a); DTM after opening and Gaussian filters (b) on a hillslope of the study area

D. LiDAR and Pleiades DEMs comparison

Because the Pleiades DSM was acquired 12 years later than the LiDAR DSM and the aim of the work is to extract the man-made structures in the landscape, the elevation of the two datasets was compared only for permanent structures (roads, ways, terraces). Due to the non-Gaussian deviation distribution with short tails, robust statistics were computed by comparing the elevations at the remaining 87000 grid nodes. The absolute median deviation between the Pleiades and LiDAR elevations was estimated to be 0.35 m. Overall, 90% of the deviations belong to [-1.4m, +1.1m], and 50% of the deviations belong to [-0.17m, +0.42m]. As pointed out by [39] the feasibility of Pleiades datasets and the quality of the derived DEMs in relation to the incidence angle depends on the context: in open landscapes without severe occlusions, the use of a single stereo pair can provide optimal results also with a fairly wide stereo angle. Giving the initial incidence angle of the considered dataset (30 degrees), despite being more noisy, the overall quality of the Pleiades derived elevation surfaces is close to that derived from the airborne DEMs. As well, despite the local scale roughness, the Pleiades DSM shows an elevation disruption for every terrace location, as shown in figure 4.

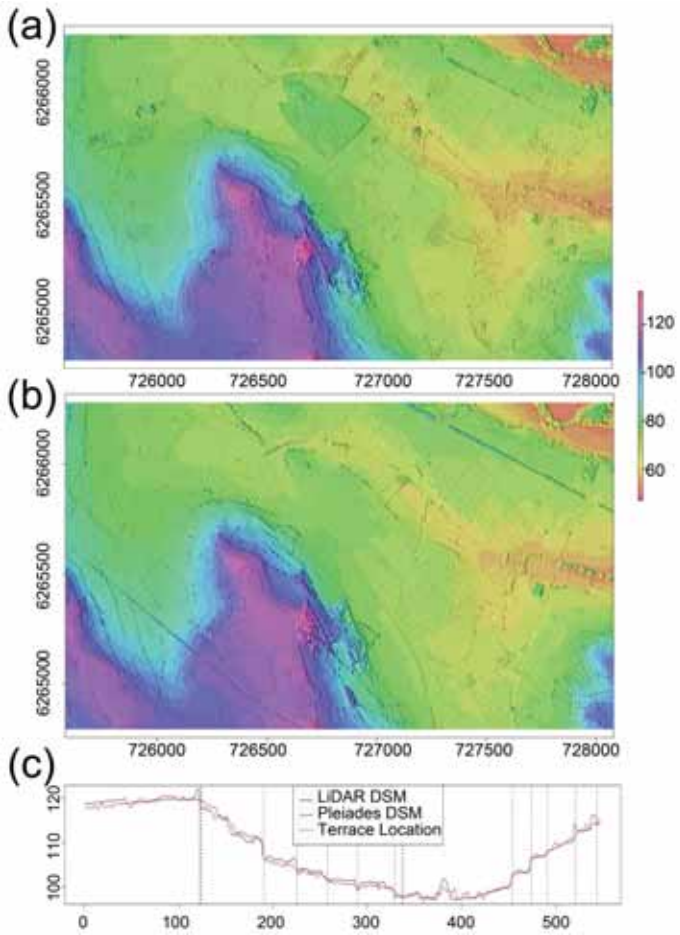


Fig. 4. Pleiades and LiDAR DSM comparison. The 2002 Lidar DSM (a) is compared to the 2013 Pleiades DSM (b). The figure also shows an elevation profile of a hill slope and the terrace locations (c)

III. Methods

A. Terrace slope detection using the Fast Line Segment Detection (LSD) image processing algorithm

This detection method relies on two steps: a DTM or DSM local contrast enhancement step providing a grey level image from which a final feature extraction step (LSD) is performed (Fig. 5).

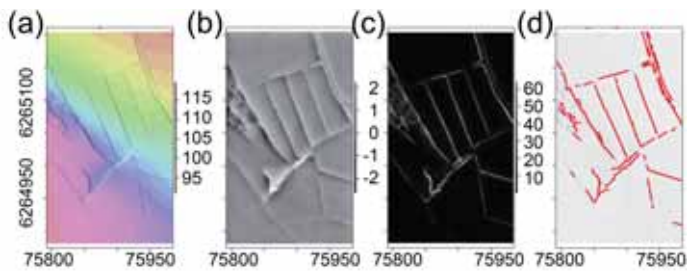


Fig. 5. LSD method: (a) original DEM, (b) DEM local contrast enhancing from a Gaussian filter, (c) Gradient magnitude image and (d) detected line segment in red

The LSD aims at detecting locally straight contours on images, with a contour being an area where the gray level is changing fast enough from dark to light or the opposite [41]. To obtain a monochromatic image and enhance terrraces as contours, the elevation contrast of the DEM was increased by subtracting a regularised DEM after Gaussian smoothing with a radius of 10 pixels to the initial DEM (Fig. 5c).

Next, the image is processed using the Fast Line Segment Detection algorithm [40]. The following paragraphs will expose its main characteristics and parameters, despite its presentation as a parameterless method [41].

Given an image gradient (Fig. 5c), the method computes a level-line field, where level-lines are created at each pixels as vectors orthogonal to the gradient direction. The level-line field is then segmented into connected regions of pixels that have the same level-line orientation, up to a certain toleration angle τ (Fig. 6). Each region is considered as a possible line candidate. The acceptation-rejection of a region as a line segment follows the Helmholtz principle, which states that no line segment should be detected in an image of pure noise.

The acceptation-rejection relies on the following processes: i) a rectangle entirely covering each region is delineated (Fig. 6); ii) the center of mass of the region is used to select the center of the rectangle, and the first inertia axis is considered to select the rectangle orientation; iii) for each rectangle, the total number of pixels in the rectangle, n, and its number of aligned points (level-lines having their orientation within the τ tollerance), k, are counted; iv) for each rectangle, the score $\frac{k}{n}$ is calculated; v) this score is compared with a binomial k distribution under an hypothesis H_0 of pure random angle fields (noise) with a distribution angle of $[0, \pi]$; vi) the region is accepted as a line segment when the score is rare in the H_0 distribution (low probability). This latter test is also referred to as a *a contrario* approach [41].

In a final step, a vectorised line segment list is produced from the rectangular properties of the image coordinates.

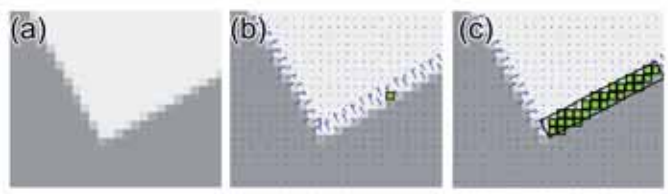


Fig. 6. Examples of a synthetic image (a) of gradient (blue) vectors, direction and magnitude (b) and a region with aligned (green) pixels on an image edge that was approximated by a rectangle (c). Images modified from [40]

Six parameters are required in the LSD algorithm: τ, S, q, e, d, c. The tolerance angle τ is the first of parameter and is expressed in degrees with a default value of 22.5 as supported by tests on numerous images [40].

The scale factor S controls the image size reduction and avoids some artifacts in the computation. S is expected to range from 0.8 (no scaling) and up to ∞. The c parameter controls the Gaussian kernel width (standard deviation) and is equal to $\frac{c}{S}$ with a default c value of 0.6.

The q parameter controls the gradient threshold $\rho = \frac{q}{sin(\tau)}$. Pixels with gradient magnitudes smaller than ρ are excluded when a region grows. q is generally fixed to 2, and should change when the image intensity range differs from the $[0, 255]$ interval.

The e parameter denotes the acceptation probability threshold of the k distribution under H_0. However the LSD algorithm is hardly sensitive to this parameter [41].

The last control parameter is the minimal density d of the region of aligned pixels in the rectangle to be accepted. This parameter is considered in addition to the *a contrario* test and is expected to range from 0 to 1. The default d value was fixed to 0.5.

The retained optimised parameter values for this study case are exposed in Table I. Only the q and e parameters were optimised; the other parameters were fixed with values of $\tau = 22.5$, $c = 0.6$, $d = 0.5$, $S = 3$. In this study case, we generally selected q=3 in order to minimize the effect of local noise (local bumps). A higher value for q was used for the LiDAR DSM. This dataset was acquired during the vines vegetative period, thus the vine rows gave high linear local gradient on the map: this high gradient was compensated by selecting the higher q. This was not needed for the Pleiades DSM, aquired at the non-vegetative period.

LSD parameters	q	e
LiDAR dtm	3	20
LiDAR dsm	8	20
Pleiades dtm	3	80
Pleiades dsm	3	80

TABLE I
LSD PARAMETER VALUES

B. Terrace slope detection using a geomorphometric algorithm

Physical processes and anthropic elements leave important topographic signatures that can be studied using distribution analysis. In addition, recent literature has underlined how statistical analyses along with LiDAR derived topographic parameters facilitate the objective recognition of different types of landscape features and processes (e.g. [42], [43]). For this study, we applied the feature extraction technique proposed by [44] and effectively tested by [4] for identifying terrace walls. According to this approach, the surface morphology is approximated by a quadratic function (Eq. 3).

$$z = ax^2 + by^2 + cxy + dx + ey + f \tag{3}$$

where x, y, z are local coordinates, and a through f are quadratic coefficients.

Any terrace wall represents a ridge on the side of the hill; therefore, the maximum curvature C_{max} (Eq. 4) can be considered as an optimal parameter. C_{max} is derived by solving and differentiating equation 3 within a local moving window as follows [45]:

$$C_{max} = kg(-a - b + \sqrt{(a-b)^2 + c^2}) \tag{4}$$

where k denotes the size of the moving window and g denotes the DEM resolution.

Anthropogenic elements such as terrace slopes present clearly defined boundaries with much sharper shapes than natural terrain features. Consequently, these elements can be identified as outliers in the positive tail of the C_{max} distribution [44]. Considering a robust statistic approach [46], these elements can be identified as outliers of a whisker plot, as the points verifying equation 5.

$$C_{max} > Q_3 + 1.5IQ \tag{5}$$

where Q_3 and IQ respectively denote the 3rd quartile and inter-quartile distance of the C_{max} distribution.

The thresholding approach (Eq. 5) produces a raster map of the potential terraces. To calculate the vector map of the terraces that correspond to the raster centerline, we use a method that is similar to the one proposed by [47]. This approach is based on the Marr-Hildreth technique [48] and is less computationally intensive than other methods [49]. The algorithm divides the raster into subsets of 100 x 100 pixels that are processed simultaneously and in parallel. For each subset, the algorithm focuses on each extracted region of pixels and processes it using the following steps. (i) Determining the Euclidean distance from each feature pixel to the nearest background pixel. (ii) The distance map is convolved with a bidirectional Laplacian filter in a manner similar to that of the Marr-Hildreth edge detection algorithm [48]. (iii) To obtain the centerline, [47] suggests using a threshold between 0.7 and 0.9 to apply to the output of the Laplacian convolution. Anything below this threshold is considered part of the centerline. The authors noted that the threshold value has little impact on the final line. However, the threshold is important for ensuring the continuity of the centerline, with a higher value resulting in a more robust centerline calculation. After the different trials in this study, we considered 0.5 as the optimal threshold. (iv) The final centerline can be more than one pixel wide; therefore, the produced output is further thinned to reduce the centerline to a width of one pixel and the final centerline is converted into a vector line.

C. Detection performance analysis

Quantitative assessments of the terrace slope detection methods were performed using the buffer method, initially proposed by [50] to quantify road detection performances. This method matches the overlapping of the detected and reference vectorised network, with performance metrics based on network lengths. In addition, this method can be applied to either a connected or unconnected network, which is the case for a terrace slope network. More specifically, this method measures the length of the detected network included (TP_1) or not included in the reference domain and the length of the reference network included (TP_2) or not included in the detected domain. The domains are defined as buffers around the two networks. The width of the buffer is the only required parameter, which was fixed to 10 m to account for the planimetric accuracies of each dataset.

First, three usual performance metrics are computed from the matching: the true positive metric (TP), which is the average between TP_1 and TP_2; the true negative TN (omission), which is the length of the reference not included in the detected network domain, and the false positive FP (commission),

which is the length of the extracted network not included in the reference domain. These metrics are normalised to the total length of the reference network to a range of [0,1].

An additive performance metric stemming from the previous metrics provide a quantitative assessment of the terrace detection, which makes an objective comparison between several detection methods possible [51]. This metric, also named quality index, QI, ranges from 0 (lowest quality) to 1 (highest quality) (Eq. 6).

$$QI = \frac{TP}{TP + TN + FP} = \frac{TP}{1 + FP} \tag{6}$$

Finally, because terrace slopes are exclusively located at the borders of agricultural fields and because plot databases are frequently available, performance metrics were computed from the raw terrace segment lists out of the detection methods and from reduced lists, excluding the terrace segments that were located far away (10 meters) from the plot border lattice. This consideration only changes the FP statistic computation. Furthermore, $FP1$ denotes the FP metric computed for reduced lists and $FP2$ denotes the FP metric computed for the raw complete list.

IV. Results

A. Detection performance results: LiDAR DEMs

The performance metrics of the terraces delineation from LiDAR DSM and DTM are exposed in table II. The results showed that at least 95% to 98% of the terrace slopes were detected on DTM, regardless of the method. The LSD method is producing fewer commissions and providing the highest quality index value. Figure 7 shows that TN occurs in the LiDAR DTM when using the LSD method on flat areas where the terraces slopes have low heights. Similarly, FP occurs on the LiDAR DTM using the LSD method on i) the steepest slopes, where bushes and hedgerows can perturb the accuracy of the DTM, and ii) along main roads and down to the streams in areas where river or ditch banks can be confused with terrace slopes. Similar behaviour can be observed when using the geomorphometric method. The results in table II show similar and high raw commissions when using the LSD and geomorphometric methods. However, commissions are highly reduced thanks to the plot border filter post-processing.

When applying the methods to the LIDAR DSM, the detection performances were reduced with true positive rates of approximately 75 % for both methods. However, the geomorphometric method slightly outperforms the LSD method when applied to LiDAR DSM. This result suggests that the geomorphometric method may be more robust when applied to noisy DEM.

Method	DEM Type	TP	TN	FP 1	FP 2	QI (1)
LSD	dtm	0.95	0.05	0.20	0.33	**0.79**
LSD	dsm	0.75	0.25	0.40	0.61	**0.54**
Curv	dtm	0.98	0.02	0.36	0.68	**0.72**
Curv	dsm	0.76	0.24	0.31	0.70	**0.58**

TABLE II
DETECTION PERFORMANCES OF THE LIDAR DATASETS

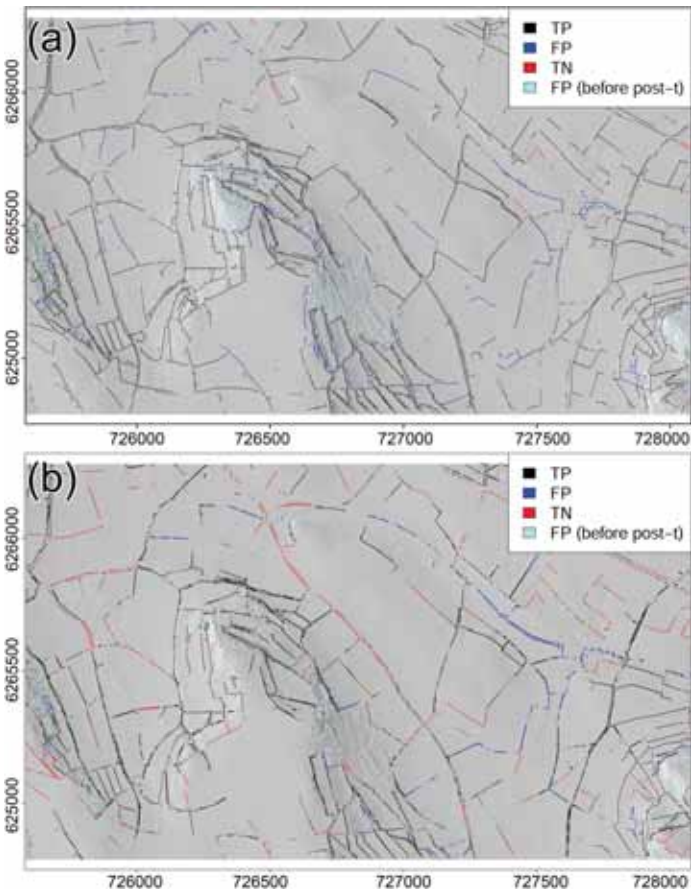


Fig. 7. Maps of the detected terraces using the LiDAR DTM and the geomorphometric method (a) or the LDS method (b). Resulting TP elements in black, TN elements in red, filtered FP elements in blue and non filtered FP elements in light blue

B. Detection performance results: Pleiades DEMs

Table III shows the performance metrics of terrace slope delineation from the Pleiades DSM and DTM.

Method	DEM Type	TP	TN	FP 1	FP 2	QI (1)
LSD	dtm	0.62	0.38	0.23	0.68	**0.50**
LSD	dsm	0.44	0.56	0.20	0.58	**0.37**
Curv	dtm	0.65	0.35	0.26	0.79	**0.52**
Curv	dsm	0.50	0.50	0.22	0.87	**0.41**

TABLE III
DETECTION PERFORMANCES OF THE PLEIADES DATASETS

The results show that only 65% of the terrace slopes are detected on the Pleiades DTM when using the geomorphometric method. The LSD method is producing slightly fewer commissions if compared with the geomorphometric method, but at the same time it provides less TP values. In general, the detection performance is systematically lower (-15%) using the Pleiades DSM. Commissions are numerous, but the results are however improved using the plot border filter post-processing. Results on the Pleiades DTM shows quite similar performances using the LSD or the geomorphometric method. The geomorphometric method, however, outperforms the LSD method when applied on the Peiades DSM: for this dataset,

the LSD seems very sensitive to DEM noise. Similarly to the LiDAR results, Figure 8 shows that TP occurs in the Pleiades DTM when using the geomorphometric method on the steepest slopes where terrace slopes are tall. FPs occur in the Pleiades DTM when using the geomorphometric method on areas where bushes and hedgerows can perturb terrace slope representation, expecially along plot borders.

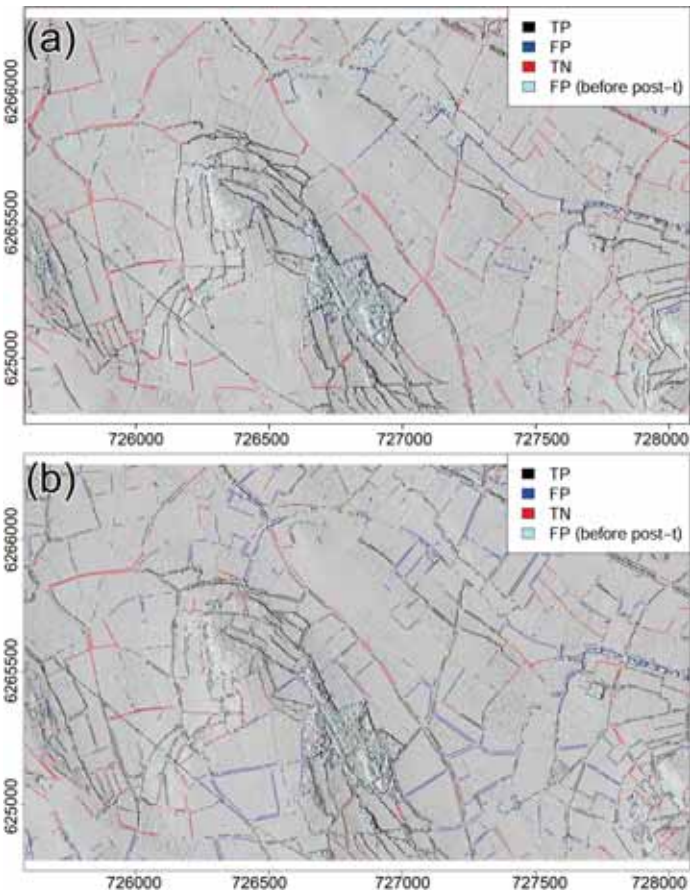


Fig. 8. Maps of the detected terraces using the Pleiades DSM and the geomorphometric method (a) or the LDS method (b).Resulting TP elements in black, TN elements in red, filtered FP elements in blue and non filtered FP elements in light blue

However, except when using the LSD method in the Pleiades DSM, the detection rate is always higher than 70 % when the terrace height is greater than 3 m (Fig. 9). The detection rate of the terrace heights is greater than 3 m and is at least 80 % when using LiDAR or the Pleiades DTM. Figure 9 also clearly shows that the LSD method fails in terrace slope detection when using the Pleiades DSM, regardless of the slope height.

V. Discussion

A. Pleiades vs LiDAR performances

The effectiveness of Pleiades for extracting anthropogenic features has been proved in other studies, mainly for building detection [27]. However, this study shows that the very high spatial resolution of the datasets allows the detection of smaller features on the earth's surface, such as terraces with heights comparable to the DEM resolution (1m). Pleiades

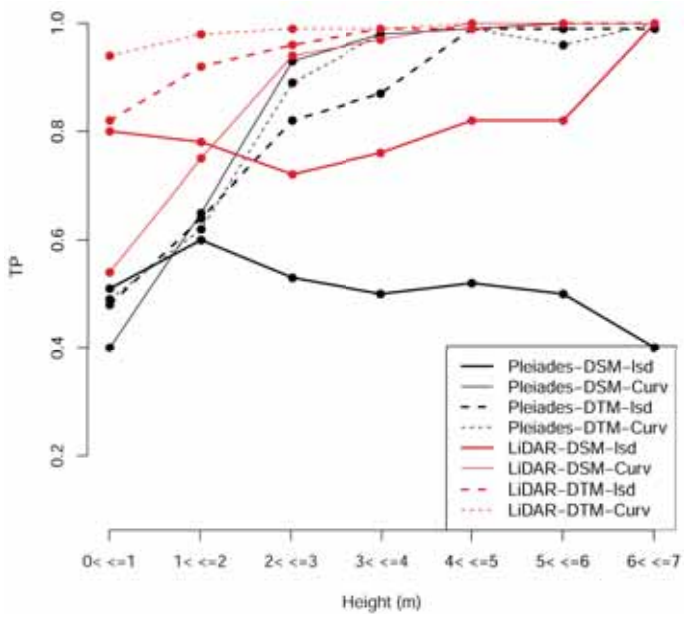


Fig. 9. Detection rate (TP) as a function of the terrace height

DTMs provide results that are comparable with the analyses performed with LiDAR DSMs. On the other hand, Pleiades DSMs have a slightly lower performance that is, however, comparable with other referenced LiDAR DTM analyses of more complex features under vegetation cover in mountainous areas [52]. This latter point can be explained by the roughness of the models. In the Pleiades DSM, being derived from satellite images, the morphology of specific features can be masked by the presence of objects on the ground as well as by their shadows. In addition, the incidence angle of the satellite might have influenced the quality of the dataset: even if the overall median difference between the LiDAR and the Pleiades datasets is low (35 cm), locally, there might be a local loss of resolution as well as some spatial artefacts induced by the incidence angle, causing the lower accuracy of the extraction mainly due to a lower completeness of the extracted features (higher number of TNs and lower number of TPs if compared to the average values of LiDAR DSMs). The size of the analysed features (height) is only slightly important when considering LiDAR DTMs and the effectiveness of detections (TPs): both the geomorphometric and the LSD algorithm quality converge when the features tend to increase in size. By contrast, when using noisy datasets (LiDAR DSM or Pleiades DEMs) the feature size is important for determining the effectiveness of the extractions, especially when using the geomorphometric approach. Specifically, when features are greater than approximately 3 m, the rate of TPs is similar for the LiDAR and Pleiades datasets. In our case, for local slope discontinuities (terraces), their vertical position in the Pleiades dataset may be underestimated due to a lower correlation scores especially with a high B/H ratio, however, the terraces are well localized in the image and this is sufficient for the proposed application: overall, the performances of the Pleiades DSMs are comparable with those of the LiDAR DSMs. However, the Pleiades system allows wider coverage

of approximately 100 x 100 km^2, which can be acquired by the satellite from the same orbit, and offers a new tool for large scales applications when compared with LiDAR DEMs.

B. Automatic mapping of terraces

Previous research regarding the extraction of features using curvature on LiDAR DTMs highlights how curvature can produce results with high branching (higher number of FP), which corresponds with greater extraction completeness [44]. Similar conclusions can be drawn when considering LiDAR DTMs and the proposed geomorphometric method, which produces a higher number of TPs and lower TN values when compared with the LSD method. When using the same topographic information, the geomorphometric method performs better when the terraces are less than 3 m tall and produces a slightly lower number of omissions. This result corresponds with the findings of [44], which underline how using curvature could be better for identifying smaller features. The use of algorithms based on the linearity of features (LSD) rather than on their convexity (geomorphometric) appears to work in the presence of more clean datasets (DTMs), independently of the considered remote sensor (LiDAR vs Pleiades). Linear detection provides a slightly higher reliability for neat datasets from LiDAR (DTMs).

C. Future challenges: DSM or DTM, filtering problems, additional required processing, dataset resolution

Considering the results in Table II and III and the performances displayed in Figure 8, the use of DTMs seems more appropriate when using LiDAR surveys and photogrammetry from Pleiades. Generally, DSMs are too noisy for comparable feature detection, but they produce good results if compared to other works in literature based on different sources and landscapes (e.g. DTMs from LiDAR [52], [53]). However, if features are large enough (generally higher than 3 m), the use of the geomorphometric approach gives good and comparable results using DSMs and DTMs. For this work the only post-processing procedure considered was a filter based on the plot border. Further improvements could be obtained, especially when using the geomorphometric approach. For example, the skeletonisation process could be improved and the looping and branching of the vector map could be reduced. In addition, using a pre-processing filter on the DSMs (median filter, or Gaussian filter, for example) could reduce the noise improving the performance of the algorithms. Further, to test the full Pleiades performances, it will be interesting to evaluate terraces extraction using the full resolution DSMs (smaller than 1 m, in some areas), to investigate the effect of noises (due to errors or surface roughness) on higher resolution DEMs (similarily to what has been done in [53]–[55]) when comparing extracted features with the ground truth datasets. These points are still open for future research.

VI. Conclusion

Three major results were obtained from this study. From the algorithm point of view, the analysis underlined that the geomorphometric algorithm is generally more robust than the LSD algorithm when applied to noisy DEMs. This information could be usefull in general also for other researches dealing with feature extractions from DEMs from multiple sources. Considering the topographic information, the automatic mapping of terraces from Pleiades DEMs is reliable for terraces that are higher than 2 m, also thanks to a post-process filter to avoid commissions far away from the plot boundaries. LiDAR detection performances are systematically higher than the Pleiades performances, especially when using a DTM. However, the filtering of the Pleiades DSMs and the successive creation of DTMs seems to give promising results, expecially considering the wider coverage that can be obtained by this dataset compared to the LiDAR coverage.

Acknowledgment

The authors would like to thank the French Space Agency CNES and Airbus Defense & Space for providing the Pleiades Images within the RCT Pleiades program. In addition, the authors would like to thank IGN-Espace, especially Mr. Nicolas Champion, for helping generate Pleiades DEMs from image pairs. Resources for the geomorphometric method computation were provided by the Interdepartmental Research Centre of Geomatics CIRGEO- at the University of Padova (Italy).

References

[1] J. Zalasiewicz, A. Smith, M. Hounslow, M. Williams, A. Gale, J. Powell, C. Waters, T. Barry, P. Bown, P. Brenchley, D. Cantrill, P. Gibbard, F. Gregory, R. Knox, J. Marshall, M. Oates, P. Rawson, P. Stone, and N. Trewin, "Are we now living in the Anthropocene.," *GSA Today*vol. 18, no. 2, 2008.

[2] J. Zalasiewicz, C. N. Waters, and M. Williams, "Human bioturbation, and the subterranean landscape of the Anthropocene," *Anthropocene*, vol. 6, pp. 3–9, 2014.

[3] A. G. Brown, S. Tooth, R. C. Chiverrell, J. Rose, D. S. G. Thomas, J. Wainwright, J. E. Bullard, V. R. Thorndycraft, R. Aalto, and P. Downs, "The Anthropocene: is there a geomorphological case?," *Earth Surface Processes and Landforms*, vol. 38, no. 4, pp. 431–434, 2013.

[4] P. Tarolli, F. Preti, and N. Romano, "Terraced landscapes: From an old best practice to a potential hazard for soil degradation due to land abandonment," *Anthropocene*, vol. 6, pp. 10–25, 2014.

[5] P. Tarolli and G. Sofia, "Human topographic signatures and derived geomorphic processes across landscapes," *Geomorphology*, 2016.

[6] A. Bevan and J. Conolly, "Terraced fields and Mediterranean landscape structure: An analytical case study from Antikythera, Greece," *Ecological Modelling*, vol. 222, pp. 1303–1314, 2011.

[7] European Commission, "The Common Agricultural Policy a Partnership between Europe and Farmers," *Directorate General Agriculture and Rural Development. Publications Office of the European Union*, 2012.

[8] T. Douglas, D. Critchley, and G. Park, "The deintensification of terraced agricultural land near trevelez, sierra nevada, spain," *Global Ecology and Biogeography Letters*, pp. 258–270, 1996.

[9] J. Lesschen, L. Cammeraat, and T. Nieman, "Erosion and terrace failure due to agricultural land abandonment in a semi-arid environment," *Earth Surface Processes and Landforms*, vol. 33, no. 10, pp. 1574–1584, 2008.

[10] J. Dover, T. Sparks, S. Clarke, K. Gobbett, and S. Glossop, "Linear features and butterflies: the importance of green lanes," *Agriculture, ecosystems & environment*, vol. 80, no. 3, pp. 227–242, 2000.

[11] J. Blondel, "The design of mediterranean landscapes: a millennial story of humans and ecological systems during the historic period," *Human Ecology*, vol. 34, no. 5, pp. 713–729, 2006.

[12] R. Z. Wheaton, *Terracing as a Best Management Practice for Controlling Erosion and Protecting Water Quality*. Purdue University, Cooperative Extension Service, 2001.

[13] D. W. Hallema and R. Moussa, "A model for distributed giuh-based flow routing on natural and anthropogenic hillslopes," *Hydrological Processes*, vol. 28, pp. 4877–4895, 2013.

[14] R. Cots-Folch, J. A. Martínez-Casasnovas, and M. C. Ramos, "Agricultural trajectories in a Mediterranean mountain region (Priorat, NE Spain) as a consequence of vineyard conversion plans," *Land Degradation & Development*, vol. 20, pp. 1–13, 2009.

[15] E. Cammeraat, R. van Beek, and A. Kooijman, "Vegetation Succession and its Consequences for Slope Stability in SE Spain," *Plant and Soil*, vol. 278, pp. 135–147, 2005.

[16] M. Koulouri and C. Giourga, "Land abandonment and slope gradient as key factors of soil erosion in Mediterranean terraced lands," *CATENA*, vol. 69, pp. 274–281, 2007.

[17] J. P. Lesschen, L. H. Cammeraat, and T. Nieman, "Erosion and terrace failure due to agricultural land abandonment in a semi-arid environment," *Earth Surface Processes and Landforms*, vol. 33, no. 10, pp. 1574–1584, 2008.

[18] P. Tarolli, G. Sofia, S. Calligaro, M. Prosdocimi, F. Preti, and G. Dalla Fontana, "Vineyards in terraced landscapes: new opportunities from lidar data," *Land Degradation and Developement*, vol. 26, pp. 92–102, 2015.

[19] J. Arnáez, N. Lana-Renault, T. Lasanta, P. Ruiz-Flaño, and J. Castroviejo, "Effects of farming terraces on hydrological and geomorphological processes. A review," *CATENA*, vol. 128, pp. 122–134, 2015.

[20] G. Sofia, F. Marinello, and P. Tarolli, "A new landscape metric for the identification of terraced sites: the slope local length of auto-correlation (sllac)," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 96, pp. 123–133, 2014.

[21] F. Gallart, P. Llorens, and J. Latron, "Studying the role of old agricultural terraces on runoff generation in a small mediterranean mountainous basin," *Journal of Hydrology*, vol. 159, no. 1-4, pp. 291 – 303, 1994.

[22] F. Levavasseur, *Structure du paysage et fonctionnement hydrologique. Application aux réseaux de fossés en zone viticole méditerranéenne.* PhD thesis, Montpellier Supagro - École Doctorale SIBAGHE, 2012.

[23] S. Price and L. Nixon, "Ancient Greek agricultural terraces: Evidence from texts and archaeological survey," *American Journal of Archaeology*, vol. 109, no. 4, pp. 665–694, 2005.

[24] T. V. D. Sluis, T. Kizos, and B. Pedroli, "Landscape Change in Mediterranean Farmlands: Impacts of Land Abandonment on Cultivation Terraces in Portofino (Italy) and Lesvos (Greece)," *Journal of Landscape Ecology*, vol. 7, pp. 23–44, 2014.

[25] J.-S. Bailly and F. Levavasseur, "Potential of linear features detection in a mediterranean landscape from 3d vhr optical data: application to terrace walls," in *IGARSS*, (Munich, Germany), pp. 7110–7113, 2012.

[26] R. Diaz-Varela, P. J. Zarco-Tejada, V. Angileri, and P. Loudjani, "Automatic identification of agricultural terraces through object-oriented analysis of very high resolution dsms and multispectral imagery obtained from an unmanned aerial vehicle," *Journal of environmental management*, vol. 134, pp. 117–126, 2014.

[27] D. Poli, F. Remondino, E. Angiuli, and G. Agugiaro, "Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pléiades-1A stereo images for 3D information extraction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 100, pp. 35–47, 2015.

[28] M. Bernard, David, Decluseau, L. Gabet, and P. Nonin, "3d capabilities of pleiades satellite," *International archives of the photogrammetry. Remote Sens Spat Inf Sci*, vol. 39, pp. 553–557, 2012.

[29] K. Jacobsen and H. Topan, "Dem generation with short base length Pleiades triplet," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-3/W2, pp. PIA15+HRIGI15, 2015.

[30] P. Tarolli, "High-resolution topography for understanding earth surface processes: Opportunities and challenges," *Geomorphology*, vol. 216, pp. 295–312, 2014.

[31] A. Baudoin, E. Boussarie, P. Damilano, G. Rum, and F. Caltagirone, "Pléiades: A multi mission and multi co-operative program," *Acta Astronautica*, vol. 51, pp. 317–327, 2002.

[32] A. Temple-Boyer, J.-F. Richard, and P. Arnould, "Segmenter les paysages de l'eau : une methode pour l'interpratation hydrodynamique des paysages (Dorsale tunisienne)," *Science et changements planétaires / Sécheresse*, vol. 18, pp. 149–160, 2007.

[33] M. Pierrot-Deseilligny and N. Paparoditis, "A multiresolution and optimization-based image matching approach: An application to surface reconstruction from spot5-hrs stereo imagery," in *In: Proc. of the ISPRS Conference Topographic Mapping From Space (With Special Emphasis on Small Satellites), ISPRS*, 2006.

[34] M. Pierrot-Deseilligny, L. De Luca, and F. Remondino, "Automated Image-Based Procedures for Accurate Artifacts 3D Modeling and Orthoimage Generation," *Geoinformatics FCE CTU*, vol. 6, pp. 291–299, 2011.

[35] A.-M. Rosu, M. Pierrot-Deseilligny, A. Delorme, R. Binet, and Y. Klinger, "Measurement of ground displacement from optical satellite image correlation using the free open-source software MicMac," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 100, pp. 48–59, 2015.

[36] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, pp. 328–41, 2008.

[37] G. Vosselman, "Slope based filtering of laser altimetry data," *International Archives of Photogrammetry and Remote Sensing*, vol. 33, no. B3/2; PART 3, pp. 935–942, 2000.

[38] J.-S. Bailly, P. Lagacherie, C. Millier, C. Puech, and P. Kosuth, "Agrarian landscapes linear features detection from lidar: application to artificial drainage networks," *International Journal of Remote Sensing*, vol. 29(12), pp. 3489–3508, 2008.

[39] M. Bernard, D. Decluseau, L.Gabet, P. Nonin, "3D capabilities of pleiades satellite," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIX, no. B3, 2012.

[40] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "Lsd: A fast line segment detector with a false detection control," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 4, pp. 722–732, 2010.

[41] R. G. Von Gioi, *A Contrario Line Segment Detection*. Springer, 2014.

[42] G. Sofia, F. Marinello, and P. Tarolli, "Metrics for quantifying anthropogenic impacts on geomorphology: road networks," *Earth Surface Processes and Landforms*, 2015.

[43] G. Sofia, P. Tarolli, F. Cazorzi, and G. Dalla Fontana, "Downstream hydraulic geometry relationships: Gathering reference reach-scale width values from LiDAR," *Geomorphology*, vol. 250, pp. 236–248, 2015.

[44] G. Sofia, G. D. Fontana, and P. Tarolli, "High-resolution topography and anthropogenic feature extraction: testing geomorphometric parameters in floodplains," *Hydrological Processes*, vol. 28, no. 4, pp. 2046–2061, 2014.

[45] J. Wood, *The geomorphological characterisation of digital elevation models*. PhD thesis, University of Leicester, 1996.

[46] J. W. Tukey, "Exploratory data analysis," 1977.

[47] T. M. Pavelsky and L. C. Smith, "RivWidth: A Software Tool for the Calculation of River Widths From Remotely Sensed Imagery," *Geoscience and Remote Sensing Letters, IEEE*, vol. 5, pp. 70–73, 2008.

[48] D. Marr and E. Hildreth, "Theory of Edge Detection," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 207, pp. 187–217, 1980.

[49] D. Ziou and S. Tabbone, "Edge Detection Techniques - An Overview," *International Journal of Pattern Recognition and Image Analysis*, vol. 8, pp. 537–559, 1998.

[50] C. Heipke, H. Mayer, C. Wiedemann, and O. Jamet, "Evaluation of automatic road extraction," *International Archives of Photogrammetry and Remote Sensing*, vol. 32, no. 3 SECT 4W2, pp. 151–160, 1997.

[51] I. Molly and T. Stepinski, "Automatic mapping of valley networks on mars," *Computers & Geosciences*, vol. 33, pp. 728–728, 2007.

[52] C.-W. Lin, C.-M. Tseng, Y.-H. Tseng, L.-Y. Fei, Y.-C. Hsieh, and P. Tarolli, "Recognition of large scale deep-seated landslides in forest areas of Taiwan using high resolution topography," *Journal of Asian Earth Sciences*, vol. 62, pp. 389–400, 2013.

[53] P. Tarolli, G. Sofia, and G. Dalla Fontana, "Geomorphic features extraction from high-resolution topography: landslide crowns and bank erosion," *Natural Hazards*, vol. 61, no. 1, pp. 65–83, 2012.

[54] F. Pirotti, P. ,Tarolli, "Suitability of LiDAR point density and derived landform curvature maps for channel network extraction," *Hydrol. Process*, vol. 24, pp. 11871197, 2010.

[55] G. Sofia, F. Pirotti, P. ,Tarolli, "Variations in multiscale curvature distribution and signatures of LiDAR DTM errors," *Earth Surface Processes and Landforms*, vol. 38, pp. 11161134, 2013.

Giulia Sofia Giulia Sofia is a Post-Doctoral Research Associate at the Department of Land, Environment, Agriculture and Forestry-University of Padova (Italy). She received a B.S. and M.S. in Forestry Science, and a Ph.D. degree (2012) in Water Resources, Soil Conservation and Watershed Management from the University of Padova (Italy). Her area of research is geomorphology and digital terrain analysis, with a special interest in feature extraction from high resolution topography. Her recent research interest concerns anthropogenic landscapes, incorporating the related human-induced processes. http://scholar.google.it/citations?user=0i-NjLMAAAAJ&hl=it, http://orcid.org/0000-0002-9769-439X

Jean-Stéphane Bailly Jean-Stéphane Bailly received the Engineer degree in agronomy, the M.S. degree in biostatistics, and the Ph.D. degree in hydrology from the University of Montpellier, France. He has been a member of the French Civil Corps of Water and Forestry Engineers since 1991. He is senior lecturer and scientist in physical geography at AgroParisTech, Montpellier, France. His research is devoted to spatial observations and parametrizations for hydrological modeling. http://www.researcherid.com/rid/C-2391-2011

Nesrine Chehata Nesrine Chehata received the Diplom. Engn. in Physics from the Ecole Nationale Suprieure de Physique de Strasbourg, 2001 and the M.S degree in Image processing from the university of Strasbourg I, 2001. She defended her PhD Thesis in computer vision and image processing in 2005 at the university of Paris 5. She worked in collaboration with the IGN (French mapping Institute) in MATIS Lab and the CNES (French Space Agency). From 2006, she has been helding a teaching position as associate professor in the EGID institute at the university of Bordeaux which than became in 2011 the ENSEGID engineering school of Environnement, Geo-resources and sustainability). From 2012 to 2013, she was a research fellow at the IRD-UMR LISAH lab (French Institute of research for development) in Tunis, Tunisia. Her research interests are pattern recognition, machine learning and feature selection applied to remote sensing data such as airborne lidar data, hyperspectral images and VHR optical satellite images for environmental applications. She is member of the IEEE and the IEEE Remote Sensing Society.

Paolo Tarolli Paolo Tarolli is Associate Professor at University of Padova (Italy). He has held Visiting Professor positions at several universities (e.g. China University of Geosciences in Beijing, National Cheng Kung University, EPFL). He is expert in digital terrain analysis, human-Earth interactions, Earth surface processes analysis through high-resolution topography, natural hazards, shallow landslides, hydro-geomorphology, geomorphic features extraction in alpine and floodplain context, anthropogenic feature extraction in anthropogenic landscapes, LiDAR and Structure from Motion applications, and GIS. In April 2015 he was nominated Science Officer of the European Geoscience Union (EGU) for the Natural Hazard Division (NH6: Remote Sensing & Hazards). Tarolli is Editor of the journal Natural Hazards and Earth System Sciences and member of the Editorial Board of Land Degradation & Development, Earth Surface Processes & Landforms, Journal of Mountain Sciences, Anthropocene, and Heliyon. He is member of the American Geophysical Union, European Geoscience Union, British Society for Geomorphology, Asia Oceania Geosciences Society (AOGS). http://www.tesaf.unipd.it/en/paolo-tarolli http://orcid.org/0000-0003-0043-5226

Florent Levavasseur Florent Levavasseur received the Engineer degree in agronomy, and the Ph.D. degree in hydrology from the University of Montpellier, France. He is scientist in agronomy at INRA, Thiverval-Grignon, France. His research is devoted to spatial modeling of cropping systems.

5.3 Axe apprentissage automatique

LIDAR DATA CLASSIFICATION USING HIERARCHICAL K-MEANS CLUSTERING

Nesrine Chehata [a,b], Nicolas David [b], Frédéric Bretar [b]

[a] Institut EGID - Université Bordeaux 3 - Equipe GHYMAC Allée Daguin 33607 Pessac-
Nesrine.Chehata@egid.u-bordeaux3.fr
[b] Institut Géographique National - laboratoire MATIS Av. Pasteur 94165 St. Mandé cedex, France-
nicolas.david@ign.fr, frederic.bretar@ign.fr

Commission III, WG III/3

KEY WORDS: Remote Sensing, LIDAR, Hierarchical Classification, DTM, Multiresolution

ABSTRACT:

This paper deals with lidar point cloud filtering and classification for modelling the Terrain and more generally for scene segmentation. In this study, we propose to use the well-known K-means clustering algorithm that filters and segments (point cloud) data. The Kmeans clustering is well adapted to lidar data processing, since different feature attributes can be used depending on the desired classes. Attributes may be geometric or textural when processing only 3D-point cloud but also spectral in case of joint use of optical images and lidar data. The algorithm is based on a fixed neighborhood size that can deal with steep relief covered by dense vegetation, mountainous area and terrains which present microrelieves. The novelty of our algorithm consists in providing a hierarchical splitting clustering to extract ground points. The number of cluster splits is used to qualify automatically the classification reliability. This point is rarely treated in previous works. Moreover landscape predictors such as slope map are used to locally refine the classification. Finally, the methodology is extended to a multiscale framework. The hierarchical clustering is processed from coarse DTM resolution to finer one. This implementation improves the algorithm robustness and ensures reliable ground estimation. Quantitative and qualitative results are presented on the ISPRS data set.

1. INTRODUCTION

Representing the Earth's topography, that is the vegetation, the true terrain, buildings as well as any human-made infrastructures from aerial remote sensors in a 3D virtual environment has been a challenging task for scientists for many years. Recent years have seen the development of airborne scanner systems which provide dense 3D point cloud of the surface topography. This massive amount of data has to be analyzed and classified to extract pertinent informations. A Digital Terrain Model (DTM) is a fundamental layer for any application in a 3D virtual environment, and as a matter of course, plays a main role when dealing with natural risk management. Several methods have been developed for filtering lidar data to generate Digital Terrain Models. Algorithms have to process large data volumes on various and complex landscapes such as urban areas Dell'Aqua et al. (2001), forest areas Kraus and Pfeifer (1998); Haugerud and Harding (2001)
or mountainous areas Wack and Stelzl (2005). Many algorithms have been implemented and tested so far, but no generic solution appeared Sithole and Vosselman (2003). Existing works on lidar data labelling can be divided into three major approaches that will be briefly detailed hereby:

1. **Morphological filters** These filters are based on a series of 3D morphological closings and openings. Robust methods against measurement errors were proposed using a dual rank filter Eckstein and Munkelt (1995). The filter parameters highly depend on the terrain slope as well as on the relevancy of laser points to belong to the terrain: last pulse is not always a true ground point, especially in presence of dense vegetation coverage. Vosselman (2000); Sithole (2001) proposed a slope based filtering. In Kraus and Pfeifer (1998), authors have proposed an iterative linear prediction scheme to remove vegetation points in forest areas. The potential of morphological filters to provide a good estimate of the ground depends on the filtering window size. A small window size leads to a fine local topography provided that there are enough true ground points within the neighborhood. On the contrary, a large window size tends to smooth the final DTM. To overcome these effects, some authors refine locally the window size of the filter Kilian et al. (1996); Bretar and Chehata (2007). Zhang et al. (2003) have used an iterative technique using progressive morphological filters by varying the window size to estimate different height thresholds in local regions. Others propose a repetitive interpolation of DTM in forest areas Filin and Pfeifer (2006); Kobler et al. (2007) to improve the algorithm robustness. The advantage of the morphological filters is the short computing time but they need an accurate a-priori knowledge about the terrain topology.
2. **Progressive TIN densification** Some points are identified as ground points and based on those, new points will be added to the ground class Sohn and Dowman (2002). In Axelsson (2000), the authors present an iterative Triangular Irregular Network generation. From a coarse triangulated surface based on the lowest points, new lidar points are integrated in a Delaunay triangulation under strong angle and distance constraints. The advantages of triangulation based methods are the short computing time and the robustness. However, the TIN surface is very sensitive to negative outliers that may shift the surface downwards.
3. **Surface model filters** These filters are based on robust interpolation of ground points Kraus and Pfeifer (1998). A coarse surface is estimated. All points are weighted by a power function of their residuals to the approximated surface. The surface converges toward points with negative residuals. In Elmqvist (2002), the ground is estimated by an active shape

model. The drawback of this approach is that it is controlled with many parameters and is sensitive to negative outliers.
In addition to the filtering process, many authors tried to organize the 3D cloud into multiple classes, using essentially unsupervised classification methods. The input data can be only 3D point cloud. Geometric or textural attributes are used. In Elmqvist et al. (2001), the height texture is the maximal local slope and the second derivative of the pixel and the 8-neighbouring pixels. Multiple echos allow the distinction between buildings and vegetation. Height texture is often processed over a regular interpolated grid. Suitable results can be obtained by Laplace Operator Maas (1999) or by local curvature Steinle and Vogtle (2001). Standard deviation of heights is also used. It can be processed in 3D or in 2.5D, over a titled plane or over a horizontal plane Tovari (2006).
In Charaniya et al. (2004), lidar intensity is used with geometric features to provide four classes: road, grass, building and vegetation. Other methods are based on a joint use of lidar data and optical images. Spectral and geometric attributes are then used. Spectral attributes are computed to qualify vegetation such as NDVI Steinle and Vogtle (2001), HNDVI Bretar (2007) and buildings Rottensteiner et al. (2007).

Existing methods provide adaptative solutions to specific areas or may be sensitive to negative outliers. All methods do not offer an automatic qualification of class reliability. In this study, we propose to use of the well-known K-means clustering in a hierarchical approach to filter lidar data. We propose a methodology that is based only on 3D point cloud. It is a point-based classification with regard to a fixed local neighborhood. It is especially dedicated to vegetated areas where ground points are sparse. It can deal with steep relief covered by dense vegetation, mountainous area and terrains which present microrelieves. Moreover, many algorithms depend on the neighborhood window size and try to adapt it provided a priori knowledge about the terrain topology. To overcome this problem, we propose a multi-scale framework that is processed in a coarse-to-fine way. It improves the algorithm robustness to the window size and provides reliable ground estimation.

The methodology is detailed in section 2. First, section 2.1 presents the management of the point cloud. Section 2.2 details the hierarchical filtering of the point cloud using K-means algorithm that provides a robust approximated surface. The number of cluster splits is used to automatically qualify the filtering reliability. They are jointly used with a local slope map to refine the ground points filtering (cf. Section 2.3). The mutliscale extension is detailed in section 2.4. Finally, quantitative and qualitative results are presented in section 4 on the ISPRS data set.

2. METHODOLOGY

2.1 Management of the point cloud geometry

Considering the DTM as a georeferenced regular gridded surface with a resolution r, the system explores the lidar point cloud following this gridded geometry. For each site s, the local 3D environment, noted V_s, is extracted. It is designed as a cylindric neighborhood centred on s and of diameter $d_s = 2*r$ ensures *50%* overlap between 3D neighborhoods and filtering regularity.

2.2 Hierarchical K-means Filtering

We propose an unsupervised filtering based on K-means clustering of the 3D-point cloud. The clustering is processed in the feature space. Mean and standard deviation of heights are used. Generally, filtering methods depend of the window size d_s. Unlike techniques that try to adapt the window size, the idea is to use a fixed window size with a hierarchical clustering of point cloud based on series of splits of the ground cluster.
The algorithm is described in figure 1. The 3D point cloud is initialized as off-ground. For each site $s(i,j)$, the cylindric neighborhood V_s is extracted. First, negative outliers are filtered. If the percentage of negative points in V_s is over ($T_{out} = 80\%$), they are considered as outliers and labelled as Non-Determined points. The filtering starts with a coarse clustering. The centroids are initialized at equal distance on Z interval. The number of initial centroids is initialized to 1. It increases iteratively while the intra-variance cluster is higher than 1m. The number of initial clusters can go up to three which can roughly correspond to ground, off-ground and low off-ground classes. n_{split} corresponds to a cluster map with DTM resolution.

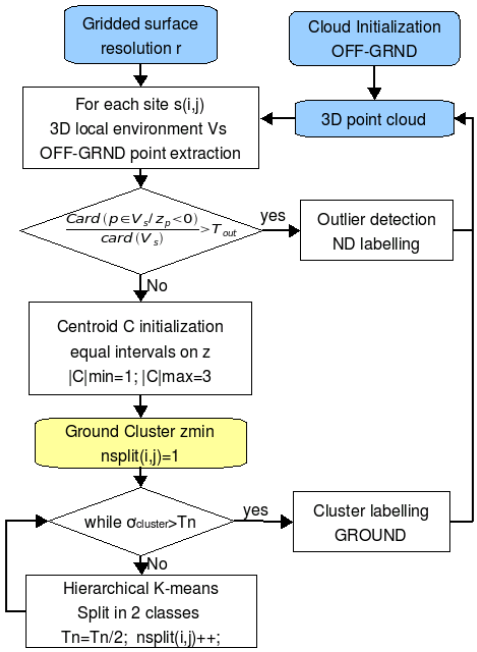


Figure 1: Hierarchical K-means flowchart.

The number of splits is stored for each site s. The cluster whose average height is minimum is considered as the ground cluster. From then, it is refined iteratively. The K-means is implemented in a hierarchical way that splits the ground cluster into two classes while the cluster standard deviation $\sigma_{cluster}$ is higher than a threshold T_n. n is the number of cluster splits. The threshold is refined at each split following $T_{n+1} = T_n/2$.
When the algorithm converges, the ground cluster is labeled and the propagation continues through the 3D point cloud. In case of DTM generation, the corresponding site is assigned to the average height of the ground cluster. This method provides a robust filtering of ground points. Therefore, 3D points labeled as ground are not reprocessed when moving to neighboring site.
The filtering step classes the point cloud into ground, off-ground and non-determined points. There is no need for pre-processing the data to filter the outliers. They are handled in the hierarchical filtering process and labeled as non-determined points. It also provides the corresponding cluster map. This map is used to qualify the classification reliability. The less the number of splits, the more reliable the classification. This

information is used to refine the classification as detailed in the following section.

2.3 Classification refinement

As cited above, our methodology is especially dedicated to vegetated and mountainous areas. The clustering process is based on point heights and tends to minimize the intra-variance of each cluster and to maximize the height difference between clusters. A neighborhood with high variance is likely to belong either to a vegetation area or to a steep surface. The most errors occur in case of steep relief with vegetation Sithole and Vosselman (2004). Therefore, sites with a high number of cluster splits (n>2) and a high local neighborhood slope (>10°) are reprocessed to refine the classification.

The estimation of the local slope is detailed in the following section. For these sites, the hierarchical K-means filtering is then processed in the fitted plane framework. The distances between clusters are unsigned to take into account both points which are above and below the estimated local plane.

Estimating the local slope Based on the assumption that the local slope changes locally in a monotonous manner, a local tangent plane is fitted, for each site s to lidar points within the cylindric neighborhood V_s. The quality of the local plane estimation depends on the lidar point distribution within V_s, the windows size and defines the terrain height relevancy as well as its uncertainty.

We estimate a plane $n_x x + n_y y + n_z z + d = 0$ with $(n_x, n_y, n_z) \in [-1, 1]$ and $d = 2 * r$. A robust M-estimator has been used with a L_p-norm, $p=1.2$ Xu and Zhang (1996). This algorithm is implemented as an iterative re-weighted least square system. At each DTM site, the steepness (the elevation angle of the surface normal) is processed, leading to a slope map. Figure 3(b) has been processed on the sample 52 of ISPRS data set. One can observe the discontinuity on the right of the image and the sharp ridges. The contribution of the local slope is discussed in Section 4.1.

2.4 Coarse to Fine implementation

The evaluation of the algorithm robustness (Section 4.1) showed that results are very sensitive to the neighboring window size and to the DTM resolution. It gives good results provided a priori knowledge. Tuning the parameters is difficult. To overcome this problem, we propose a multiscale approach for extracting ground points. The idea is to start at coarse resolution with a high window size to ensure to have some ground points. This may lead to the surface overfitting. The result is refined later at fine resolution with a small window size. The clustering method is still hierarchical.

The difference is that the fine ground cluster tends to minimize the mean and the standard deviation of differences between ground points heights and the estimated elevation (at coarse resolution). Evaluations (Section 4.3) proof that the coarse-to-fine method ensures a high reliability for ground points, that is independent of the neighborhood size. Moreover, since the finer stage is based on the coarse DTM, the methodology saves computation time.

3. THE DATA SET

The algorithm has been tested on various data sets, especially on vegetated areas with different topography. The algorithm is first analyzed on one data set and then quantitative and qualitative evaluations are presented over several others. The datasets are those made available for the ISPRSWG3/III test Sithole and Vosselman (2004). These data are considered as a reference for evaluating filtering algorithms. A ground truth of ground/off-ground classification is provided. It was performed manually in a controlled manner. In this study, we are interested on vegetated, mountainous and steep relieves. Consequently, we selected the FSite5 data set with *2m* resolution. Provided the ground truth, the incorrect classification is qualified by two measures: Type I error (classify ground points as non-ground) and Type II errors (classify non-ground points as ground). Unlike the ISPRS study, where some lidar points are unused, all the lidar points are taken into account for evaluations.

4. RESULTS AND DISCUSSION

Table 1 summarizes the characteristics of the proposed method.

Description	I/O Format	# of operator settings
Hierarchical iterative K-means	Point list Grid	3 DTM resolution neighborhood size T0 variance intra-cluster Tn

Table 1: Characteristics of the proposed method.

Compared to other methods, the algorithm depends on few parameters. The robustness of the algorithm to these parameters is evaluated in section 4.1. Results and evaluations are first detailed on sample 52 dataset. The contribution of local slope is also demonstrated. The results on previously described datasets will be compared to other filtering algorithms Sithole and Vosselman (2004). These evaluations will entail a coarse-to-fine implementation that makes the algorithm more robust and less sensible to the window size. Results will be also presented on the same areas.

4.1 Hierarchical K-means filtering

Sample 52 deals with quarry and low vegetation on river bank with gaps. The 3D point cloud is shown in figure 2.

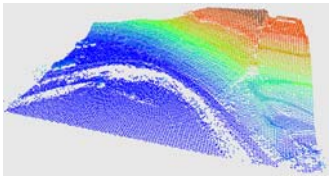
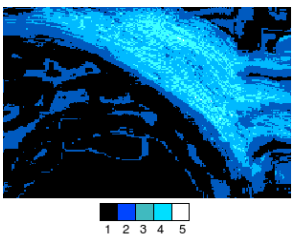
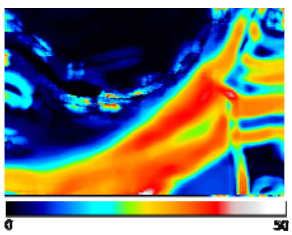


Figure 2: Sample 52-3D point cloud

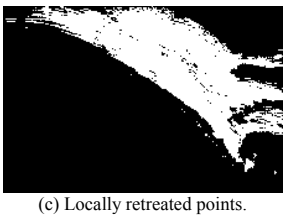
Figure 3 shows the used landscape predictors to locally refine the classification.



(a) Cluster map.



(b) Slope degree map.



(c) Locally retreated points.

Figure 3: Sample 52 - landscape predictors.

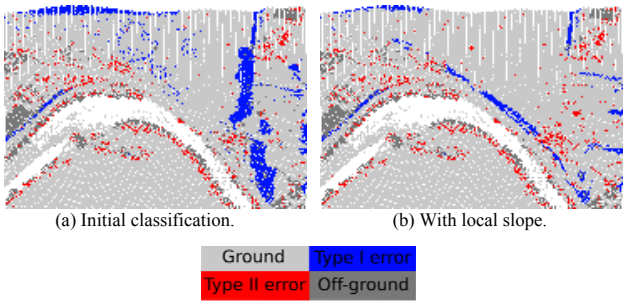
The algorithm was processed with a DTM resolution of *2m* and a neighborhood size of *30m*. The cluster split map is shown in figure 3(a). The less the number of splits, the more reliable is the ground cluster. One can observe that the number of clusters increases in the steep surface. Figure 3(b) shows the slope degree map. The combination of both maps leads to the sites that have to be reprocessed to refine the classification (figure 3(c)).

Contribution of local slope Table 2 compares the type I and type II errors after processing the confusion matrix.

Method	Type I	Type II	Total
Initial classification	7.36%	3.87%	11.23%
Classification refinement	3.10%	4.57%	7.67%

Table 2: Contribution of cluster and slope maps, DTM resolution=*2m*, neighborhood size= *30m*.

The refinement of classification decreases the total error. The type I error is improved which means that less ground points are classified as off-ground thanks to the estimation of the surface slope. Figure 4 shows the confusion images with the initial classification and the refined one using the slope map. The result is improved on the right crest where the points are classified as ground in the surface plane framework.



(a) Initial classification. (b) With local slope.

Figure 4: Sample 52 - Impact of the local slope: comparison of confusion images.

Algorithm robustness to parameters In this section, quantitative evaluations measure the algorithm robustness to its parameters. Type I, Type II and the total errors are processed. Figure 5 illustrates the impact of the neighborhood size on the evaluation with a fixed DTM resolution of *2m*. Best results are obtained with a neighborhood size $d_s = 5m$. However, when d_s increases, the type II errors are almost constant but type I errors increase. This means that more ground points are classified as non-ground. In fact, with a large d_s, the estimated plane may be erroneous in case of a relief changes and the terrain can be overfitted. Figure 6 shows the evolution of errors with the DTM resolution. The total error increases at coarse resolution of the DTM. However, the more r increases, the more type II error decreases. This property will be used in the coarse-to-fine implementation 2.4.

These evaluations show that the algorithm is sensitive to parameter tuning. For such a terrain, with a fixed neighborhood size, we still need an a-priori knowledge to choose the window size.

4.2 Comparison with filtering algorithms

In this section, table 3 shows quantitative comparison of the proposed method with other filtering algorithms described in

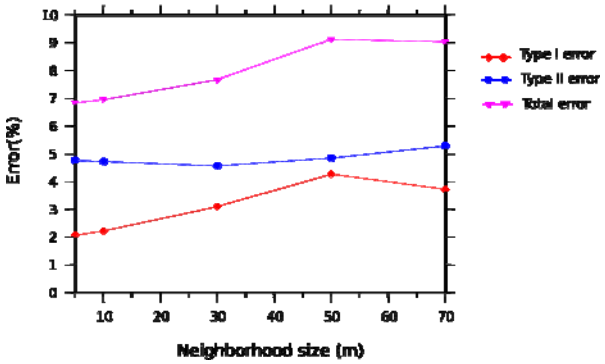


Figure 5: Evaluation of the neighborhood size impact, DTM resolution=*2m*.

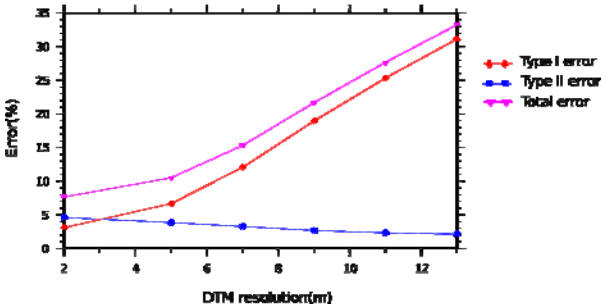


Figure 6: Evaluation of the DTM resolution impact, neighborhood size=*30m*

Sithole and Vosselman (2004).

Sample	Method	Type I	Type II	Total
Sample 51 17845 pts	Elmqvist	49,34%	1.60%	50.94%
	Sohn	10.33%	5.68%	16.01%
	Axelsson	0.13%	12.00%	12.13%
	Pfeifer	4.21%	1.93%	6.14%
	Brovelli	28.23%	3.64%	31.87%
	Roggero	1.90%	6.96%	8.86%
	Wack	14.03%	2.23%	16.26%
	Sithole	7.03%	6.99%	14.02%
	Proposed meth	0.04%	7.31%	7.35%
Sample 51 17845 pts	Elmqvist	85.05%	1.27%	86.32%
	Sohn	12.34%	9.48%	21.82%
	Axelsson	1.78%	14.21%	15.99%
	Pfeifer	21.27%	5.68%	26.95%
	Brovelli	50.43%	3.84%	54.27%
	Roggero	9.80%	9.66%	19.46%
	Wack	26.49%	1.04%	27.53%
	Sithole	30.41%	3.57%	33.98%
	Proposed meth	2.23%	4.73%	6.96%
Sample 51 17845 pts	Elmqvist	92.45%	0.18%	92.63%
	Sohn	20.48%	13.24%	33.72%
	Axelsson	8.58%	16.76%	25.34%
	Pfeifer	12.53%	14.23%	26.76%
	Brovelli	54.93%	1.62%	56.55%
	Roggero	17.81%	4.74%	22.55%
	Wack	28.33%	1.02%	29.35%
	Sithole	38.41%	4.81%	43.22%
	Proposed meth	2.15%	2.16%	4.31%

Table 3: ISPRS data sets. Quantitative evaluations of errors with comparison to the ground truth.

Three datasets of Fsite5 are used. The sample 52 which was used previously. The sample 51 is characterized by vegetation on slope. Finally the sample 53 raises the discontinuity preservation problem. For this comparison, we used a DTM resolution of *2m* and a neighborhood size of *10m*. For each sample, we highlighted the best results among filtering algorithms and the proposed method. In all cases the proposed method decreased the total errors. Moreover, it provides smaller type I errors, which means that few ground points are classified as off-ground. However, it might present heavy type II errors. This is due to the use of a relatively small neighborhood size. Some off-ground objects may be classified as ground. Figure 7 shows confusion images on samples 51 and 53. One can observe that heavy type II errors occur in sample 51 on rooftops since their size is smaller than the used neighborhood size.

4.3 Coarse to Fine implementation

For terrain modeling, the ground estimation should be robust and type II error should be minimized. Based on parameter evaluation (section 4.1), we propose a coarse-to-fine implementation to decrease type II errors. A coarse DTM resolution ensures a reliable ground estimation, the finer resolution should decrease the type I error. Figure 8 shows the evolution of confusion images from the coarse resolution to the finer one. At coarse resolution, $r=15m$ and neighborhood size $d_s = 30m$. At fine resolution $r=2m$ and $d_s=5m$. Figure 8(a) shows, at coarse resolution, heavy type I errors due to the large DTM resolution. Figure 8(b) shows the results at fine resolution. The total errors are clearly decreased. Figure 8(c) shows the spatial differences between the coarse DTM height and the fine ground cluster mean height. Negative differences are along the steep surface which was overfitted at the coarse stage. Besides, high positive differences occur on the right of the image, in presence of discontinuity.

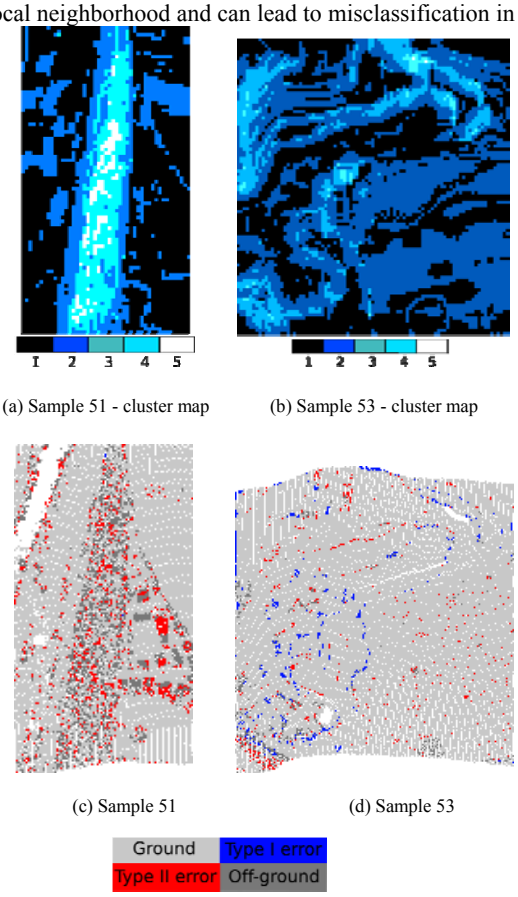
Table 4 compares the errors between the fixed neighborhood size and the coarse-to-fine implementations. Type II errors decreases with the coarse-to-fine implementation. The ground estimation is more reliable for a terrain modeling application. The coarse-to-fine implementation makes the algorithm robust the neighborhood size and the DTM resolution. It combines the advantages at each resolution. It allows to preserve the discontinuities and to deal with steep relieves.

Sample	Method	Type I	Type II	Total
51	Coarse fixed ds	12.32%	4.79%	17.11%
	Fine fixed ds	0.039%	7.31%	7.35%
	Coarse- to-fine	0.65%	6.33%	6.98%
52	Coarse fixed ds	36.99%	1.75%	38.74%
	Fine fixed ds	2.23%	4.73%	6.96%
	Coarse- to-fine	4.37%	2.99%	7.36%
53	Coarse fixed ds	34.72%	0.91%	35.63%
	Fine fixed ds	2.15%	2.16%	4.31%
	Coarse- to-fine	5.93%	1.27%	7.20%

Table 4: Contribution of cluster and slope maps, DTM resolution=*2m*, neighborhood size= *30m*.

5. CONCLUSION AND PERSPECTIVES

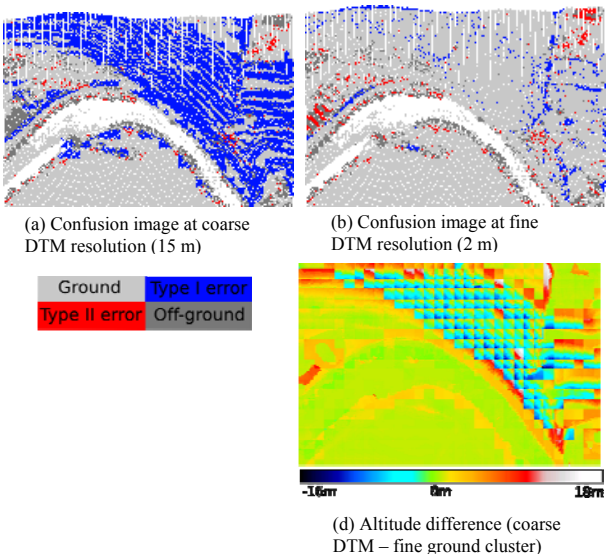
We proposed in this study, the use of K-means clustering in a hierarchical way to filter lidar data. The methodology provides good results with comparison to other filtering methods. Moreover, this algorithm provides an automatic classification quality thanks to the cluster map which is useful in case of human operator correction. The proposed classification is based on local neighborhood and can lead to misclassification in the



(a) Sample 51 - cluster map (b) Sample 53 - cluster map

(c) Sample 51 (d) Sample 53

Figure 7: Fixed neighborhood size *10m*. Confusion images on Fsite5 dataset.



(a) Confusion image at coarse DTM resolution (15 m)

(b) Confusion image at fine DTM resolution (2 m)

(d) Altitude difference (coarse DTM – fine ground cluster)

Figure 8: Sample 52 - Coarse-to-fine Approach

presence of complex objects. Landscape predictors are clearly needed to tune the parameters. We used the cluster map and a slope map to refine locally the classification. However, the parameter evaluations showed that the algorithm is sensitive to the window size and the DTM resolution. To overcome this problem, a coarse-to-fine implementation is proposed. The classification at a coarse level helps dealing with discontinuities,

gaps and complex objects. The finer step takes into account local features as slope. The ground estimation is more reliable with the mutiscale approach. The advantage of K-means clustering is that it can be easily adapted to available data by modifying feature attributes. Clustering may be then processed in a multidimensional feature space. In the proposed algorithm, only geometric attributes are used to separate ground from off-ground points. Spectral attributes could be used for fine classification purposes of off-ground objects. As a perspective, the cluster split map could be also used for a multiple resolution process of the lidar point cloud. In case of one cluster (ground), the lidar data can be under-sampled to reduce computing time.

REFERENCES

Axelsson, P., 2000. DEM generation from laser scanner data using adaptative tin models. In: International Archives of Photogrammetry and Remote Sensing, Vol. 33, part B4/1, ISPRS, pp. 110–117.

Bretar, F., 2007. Processing fine digital terrain models by markovian regularization from 3D airborne lidar data. In: IEEE ICIP, Vol. 4, pp. 125–128.

Bretar, F. and Chehata, N., 2007. Digital terrain model on vegetated areas: Joint use of airborne lidar data and optical images. In: Proc. of Photogrammetric Image Analysis (PIA), ISPRS, Munchen, Germany.

Charaniya, A., Manduchi, R. and Lodha, S., 2004. Supervised parametric classification of aerial lidar data. In: Real-Time 3D Sensors and their use, workshop in conjunction with IEEE CVPR, p. 8.

Dell'Aqua, F., Gamba, P. and Mainardi, A., 2001. Digital terrain models in dense urban areas. In: Proc. of the ISPRSWorkshop on land surface mapping and characterization using laser altimetry, IAPRS, Vol. 34, part 3/W4, Annapolis, U.S., pp. 195–202.

Eckstein, W. and Munkelt, O., 1995. Extracting objects from digital terrain models. In: Proc. Int. Society for Optical Engineering: Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes, Vol. 2572, pp. 43–51.

Elmqvist, M., 2002. Ground surface estimation from airborne laser scanner data using active shape models. In: Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision, Vol. 34, part 3A, Graz, Austria, pp. 114–118.

Elmqvist, M., Jungert, E., Lantz, F., Persson, A. and Soderman, U., 2001. Terrain modelling and analysis using laser scanner data. In: International Archives of Photogrammetry and Remote Sensing, Vol. 34, part 3/W4, ISPRS, pp. 219–226.

Filin, S. and Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS Journal of Photogrammetry and Remote Sensing 60, pp. 71–80.

Haugerud, R. and Harding, D., 2001. Some algorithms for virtual deforestation of lidar topographic survey data. In: Proc. Of the ISPRSWorkshop on land surface mapping and characterization using laser altimetry, IAPRS, Vol. 34, Annapolis, U.S., pp. 211–218.

Kilian, J., Haala, N. and Englich, M., 1996. Capture and evaluation of airborne laser scanner data. In: International Archives of Photogrammetry and Remote Sensing, Vol. 31, ISPRS, pp. 383–388.

Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Ostir, K. and Dzeroski, S., 2007. Repetitive interpolation: A robust algorithm for DTM generation from aerial laser scanner data in forested terrain. Remote Sensing of Environment 108(1), pp. 9–23.

Kraus, K. and Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53, pp. 193–203.

Maas, H., 1999. The potential of height texture measures for the segmentation of airborne laserscanner data. In: Canadian Symposium on Remote Sensing, Ottawa, Canada.

Rottensteiner, F., Trinder, J., Clode, S. and Kubik, K., 2007. Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis. ISPRS Journal of Photogrammetry and Remote Sensing 62(2), pp. 135–149.

Sithole, G., 2001. Filtering of laser altimetry data using a slope adaptative filter. In: Proc. of the ISPRS Workshop on land surface mapping and characterization using laser altimetry,Vol. 34.

Sithole, G. and Vosselman, G., 2003. Comparison of filtering algorithms. In: Proc. of the ISPRS Workshop III/3 '3D Reconstruction from Airborne Laserscanner and InSAR' – International Archives of Photogrammetry and Remote Sensing, Vol. 34, part 3/W13, ISPRS, Dresden, Germany, pp. 71–78.

Sithole, G. and Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing 59(1-2), pp. 85–101.

Sohn, G. and Dowman, I., 2002. Terrain surface reconstruction by use of tetrahedron model with the MDL criterion. In: R. Kalliany and F. Leberl (eds), Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision, Graz, Austria, pp. 336–244.

Steinle, E. and Vogtle, T., 2001. Automated extraction and reconstruction of buildings in laserscanning data for disaster management. In: Proc. of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona, Switzerland.

Tovari, D., 2006. Segmentation Based Classification of Airborne Laser Scanner Data. PhD thesis, Universitat Fridericiana zu Karlsruhe (TH).

Vosselman, G., 2000. Slope based filtering of laser altimetry data. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 33, part B3/2, ISPRS, Amsterdam, The Nederlands, pp. 935–942.

Wack, R. and Stelzl, H., 2005. Laser DTM generation for southtyrol and 3D-visualization. In: Proc. of the ISPRS Laserscanning 2005, Vol. 36, part 3/W19, Enschede, the Netherlands, pp. 48–53.

Xu, G. and Zhang, Z., 1996. Epipolar Geometry in stereo, motion and object recognition. Kluwer Academic Publishers.

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M., Yan, J. and Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne lidar data. IEEE Transactions on Geoscience and Remote Sensing 41(4), pp. 872–882.

Terrain Modeling From Lidar Range Data in Natural Landscapes: A Predictive and Bayesian Framework

Frédéric Bretar, *Member, IEEE*, and Nesrine Chehata, *Member, IEEE*

***Abstract*—The Earth's topography, including vegetation and human-made features, reduced to a virtual 3-D representation is a key geographic layer for any extended development or risk management project. Processed from multiple aerial images or from airborne lidar systems, the 3-D topography is first represented as a point cloud. This paper deals with the generation of digital terrain models (DTMs) in natural landscapes. We present a global methodology for estimating the terrain height by deriving a predictive filter paradigm. Under the assumption that the terrain topography (elevation and slope) is regular in a neighboring system, a predictive filter combines linearly the predicted topographic values and the effective measured values. In this paper, such a filter is applied to 3-D lidar data which are known to be of high elevation accuracy. The algorithm generates an adaptive local geometry wherein the elevation distribution of the point cloud is analyzed. Since local terrain elevations depend on the local slope, a predictive filter is first applied on the slopes and then on the terrain elevations. The algorithm propagates through the point cloud following specific rules in order to optimize the probability of computing areas containing terrain points. Considered as an initial surface, the previous DTM is finally regularized in a Bayesian framework. Our approach is based on the definition of an energy function that manages the evolution of a terrain surface. The energy is designed as a compromise between a data attraction term and a regularization term. The minimum of this energy corresponds to the final terrain surface. The methodology is discussed, and some conclusive results are presented on vegetated mountainous areas.**

***Index Terms*—Altimetry, digital terrain model (DTM), energy minimization, Kalman filter, lidar.**

I. Introduction

REPRESENTING the Earth's topography, i.e., the vegetation, the true terrain, and buildings, as well as any human-made infrastructures, from aerial sensors in a 3-D virtual environment has been a challenging task for scientists for many years. A digital terrain model (DTM) is a fundamental layer for any application in a 3-D virtual environment and, as a matter of course, plays a main role when dealing with natural risk management. The 3-D cartography of flooding zones, the 3-D simulation of polluted fluids, or any electromagnetic propagation fields are as much as relevant applications which rely on the accurate geometry of virtual scenes in general and particularly on the quality of a DTM.

In the context of aerial survey, images acquired in a multistereoscopic configuration can be used to derive the 3-D information of the surface using correlation techniques, such as the maximum-flow formulation [1], implemented in [2] and [3]. However, these techniques fail when radiometric inhomogeneities appear between stereoscopic couples or when textural information is not discriminant enough. Furthermore, the terrain cannot be directly estimated when it is occluded by vegetation. Airborne laser systems may sort out these problems under certain conditions. They are based on the recording of the time-of-flight distance between an emitted laser pulse and its response after a reflection on the ground. They provide sets of tridimensional irregularly distributed points, georeferenced with an integrated GPS/INS system within an elevation precision that varies between 0.05 and 0.2 m [4]–[6]. Moreover, these systems can provide multiple returns for a single laser pulse which correspond to different encountered obstacles. It is particularly relevant when surveying vegetated areas, where the altitude of the canopy, intermediate points, and even the ground can be recorded accurately with a single laser measurement, although dense vegetation can inhibit a laser ground return signal.

The aim of this paper is to present a full methodology for generating DTMs from airborne 3-D lidar data. We focused the methodology on natural landscapes that are composed of vegetated and bare Earth landscapes. Urban areas are not studied.

Once the DTM is calculated, lidar points near the terrain surface can be classified as ground points. Sithole and Vosselman [7] performed a detailed review and qualification of classification algorithms generally used when processing lidar data. Three main approaches have been developed in the literature and are briefly described hereafter.

1) The *morphological* approach is based on a series of morphological openings and closings on the 3-D point cloud [8]. A dual-rank filter has also been used to manage DTM generation [9], with an evolution of the structural element size [10].
2) The *iterative triangular irregular network generation* approach consists, from a coarse triangulated surface based on the lowest points, in integrating new lidar points in a Delaunay triangulation under strong angle and distance constraints [11].
3) The *modified kriging* approach or linear prediction [12], [13] consists in modifying the interpolator parameters depending on a weighting function associated to the lidar points. Weights are a combination of variance–covariance values which are modeled as a power function of the

Manuscript received November 19, 2008; revised June 12, 2009 and August 28, 2009. Current version published February 24, 2010.

F. Bretar is with the Laboratory of Méthodes d'Analyses pour le Traitement d'Images et la Stéréorestitution (MATIS), Institut Géographique National, 94165 St-Mandé, France (e-mail: frederic.bretar@ieee.org).

N. Chehata is with the EGID Institute, University of Bordeaux 3, 33607 Pessac, France, and also with the Laboratory of Méthodes d'Analyses pour le Traitement d'Images et la Stéréorestitution (MATIS), Institut Géographique National, 94165 St-Mandé, France (e-mail: nesrine.chehata@egid.u-bordeaux3.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2009.2032653

0196-2892/$26.00 © 2009 IEEE

residuals between the interpolated terrain value and the elevation of lidar points. The surface converges toward the points with negative residuals.

Even if these different approaches already proved to be relevant, they suffer from some drawbacks. The morphological one is certainly the fastest methodology but remains unsatisfactory over various landscapes with a unique parameter set. Depending on the window size, large objects may remain in the final DTM. The iterative triangulation has been implemented in the Terrasolid software [14], which can be used for mass DTM production. The strength of this implementation consists in the manual refinement of the terrain surface if the basic parameter set is not adequate with the entire survey. This situation may occur in the case of vegetation and steep slopes: When constraints are applied in order to avoid the vegetation to belong to the terrain surface, points located on the vegetation-free steep slopes are not included in the DTM. The modified kriging is an iterative approach which can take some time to converge and is sensitive to the input parameters.

These approaches are based on important concepts that have been used independently so far. First, the lowest lidar points in a defined neighborhood have a higher probability of belonging to the terrain than the others. Second, since the terrain surface may have been occluded either by the vegetation or by buildings, it appears to be necessary to adapt the size of the local neighborhood in order to ensure the terrain points to belong to it. Finally, most of the terrain surfaces are locally smooth (i.e., the terrain is locally autocorrelated). Our approach consists in integrating these concepts in a single methodology. Moreover, in the first part of the algorithm, we used a well-known mathematical framework (the predictive Kalman filter) which has never been used in such application. It allows modeling the terrain estimate problem, including the uncertainty propagation, independently of the landscape type.

In our implementation, a DTM will refer to an image representation of the terrain.

Section II is dedicated to the point cloud management. We describe the topology (Section II-A) and the propagation rules within the data sets (Section II-B), as well as the adaptive neighborhood strategy (Section II-C). We introduce in Section III-A the general framework of the predictive filter paradigm for DTM estimation. The integration of the local slope in this framework is highlighted and discussed in Section III-B. We then propose to apply the paradigm for calculating the terrain height in Section III-C. Since our algorithm tends to smooth the true terrain surface and since lidar points are meant to describe relevant microrelief, we introduced in our system a regularization step (Section IV) which integrates lidar points near the terrain surface. Considered as an initial surface, the DTM is regularized in a Bayesian framework. Finally, the methodology is discussed in view of some results (Section V).

The global methodology is reminded in a workflow in Fig. 1.

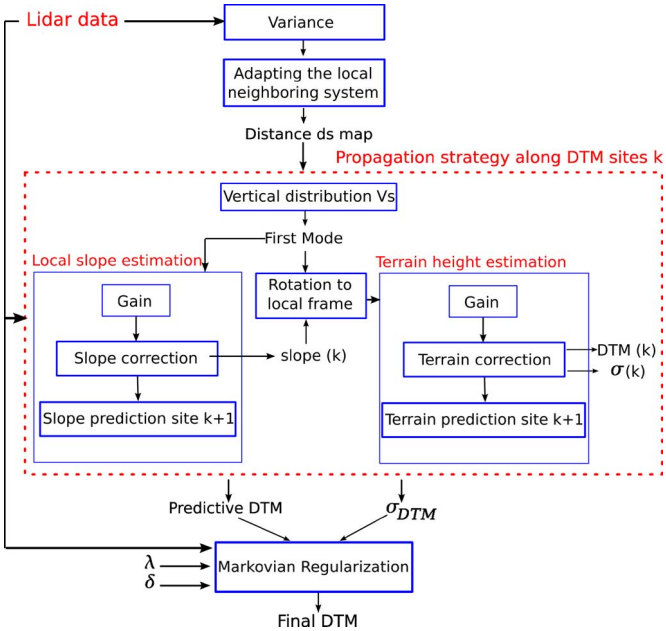


Fig. 1. Workflow of the whole methodology.

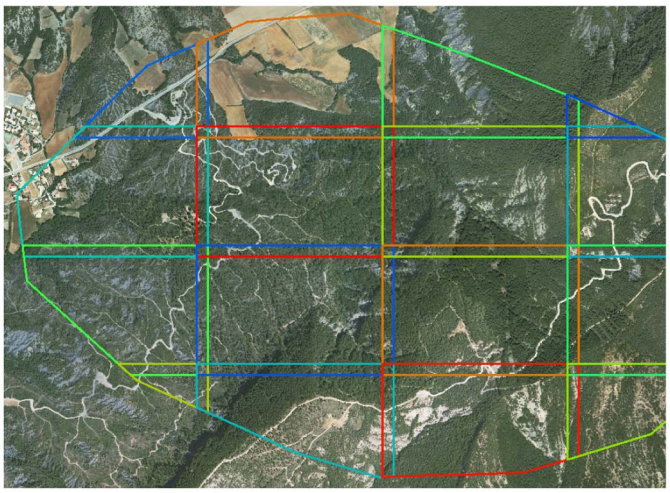


Fig. 2. Display of a paved lidar survey over a mountainous area. The background image is an orthorectified image from IGN.

II. Preprocessing and Data Management

A. Topology

A lidar survey consists of numerous binary files, either representing the entire strips or subsets of numerous strips as a paving decomposition (Fig. 2). Indeed, a lidar survey is generally composed of several millions of 3-D points. It depends on the area of the survey and on the point density. A subset of lidar points is therefore represented as sets of binary files containing the x, y, and z geographic coordinates and sometimes the lidar intensity in the near-infrared domain. After the processing, other information can be added such as a classification label (e.g., ground, off-ground, vegetation, etc.) or the normal vector direction. The point cloud topology is managed with a quadtree structure. In the following, requests on the quadtree are cylinder neighboring extractions of various diameters $\mathbf{d}_s$. We denote $\mathcal{V}_s$ such neighborhood containing the set of lidar points $l(l_x, l_y, l_z)$, s representing the corresponding image site (pixel) on the future DTM, and (s_x, s_y) the related coordinates in the 3-D space

$$\mathcal{V}_s = \{l \in \text{point cloud}/(s_x - l_x)^2 + (s_y - l_y)^2 \leq (\mathbf{d}_s/2)^2, \mathbf{d}_s \in \mathbb{R}^+\}. \quad (1)$$

B. Propagation Rules Within the Point Cloud

The trajectory of the filter within the point cloud follows specific rules: It is designed to favor the processing of sites containing terrain points before those containing only off-ground points so that the coherence of the terrain measurements should be optimized. More precisely, at a position s, the filter can move toward eight potential sites (eight-connectivity). The eligible site is chosen among the not previously visited ones and must minimize the elevation variance (denoted $[\sigma_s]_{\text{lowest}}$ in this paper) of a subset of lidar points included in $\mathcal{V}_s$. In this study, the subset was constituted by the 20% lowest points in $\mathcal{V}_s$. However, if the echo information is available, the elevation variance can be calculated on the lowest last and single echoes. This information was not provided with our data set, but we consider that a threshold of 20% over the point elevations is a fairly good approximate of the lowest last and single echoes. Here, the diameter $\mathbf{d}_s$ of $\mathcal{V}_s$ is set after a sequential search: From an initial diameter of 1 m, $\mathbf{d}_s$ is increased by the DTM resolution cell (1 m here) as long as $\mathcal{V}_s$ contains at least 25 lidar points. Moreover, it ensures that the variance calculation can be performed on at least five values. In this section, the elevation variance $[\sigma_s]_{\text{lowest}}$ has to be computed with the smallest $\mathbf{d}_s$ with regard to the point density in order to provide relevant *a priori* information whether the lidar points have hit the ground.

C. Adapting the Local Neighboring System

Depending both on the geometrical and optical properties of topographic objects, a lidar pulse may or may not penetrate off-ground objects. In the case of opaque human-made features like buildings, only one return pulse corresponds to an emitted lidar pulse. However, in the case of vegetated areas, an incident lidar pulse will interact with the canopy and be returned as a complex shape. It will be seen as multiple returns by the sensor. Mainly depending on the leaf density, a lidar measurement will be composed of several echoes: The last recorded one (farthest from the sensor) may describe the true terrain.

Processing a DTM from lidar data consists in recognizing 3-D points that belong to the terrain from those which do not. In our approach, it comes to explore the 3-D local environment $(\mathcal{V}_s)$ of each DTM site and to calculate the terrain height from the lidar point cloud (cf., Section III-A). In the case of bare Earth landscapes, the smaller the diameter of $\mathcal{V}_s$, the more realistic the estimation of the ground: There is a limited smoothing effect. However, when processing natural landscapes with vegetation, the higher the elevation variance $[\sigma_s]_{\text{lowest}}$ of the lidar points, the weaker the probability for lower points to belong to the true terrain. As a result, increasing the neighboring size $\mathbf{d}_s$, as a function of the elevation variance $[\sigma_s]_{\text{lowest}}$, will increase the probability of finding ground lidar points in $\mathcal{V}_s$.

We define an off-ground mask $\mathcal{M}$ as a binary image of the DTM like resolution. Assuming that the standard deviation $[\sigma_s]_{\text{lowest}}$ of the lidar points belonging to $\mathcal{V}_s$ is greater in the case of off-ground features than that of bare Earth areas, we consider a site s to belong to $\mathcal{M}$ if $[\sigma_s]_{\text{lowest}}$ is greater than a threshold $\sigma_{\min}$. This threshold is often set to 1 m.

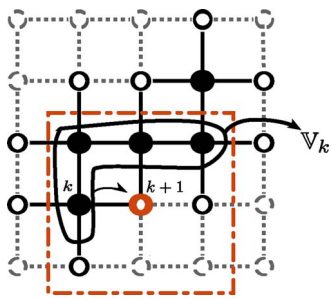


Fig. 3. Definition of $\mathbb{V}_k$. The black sites have already been processed, whereas site $k+1$ is the next one to be processed.

The neighboring size over all lidar points is initialized to a constant value $\mathbf{d}_s^{\min}$. In our case, $\mathbf{d}_s^{\min} = 5$ m. The adaptive neighboring d_s is then processed over lidar points belonging to $\mathcal{M}$. Low standard deviations of altitudes are over represented in rural areas. Therefore, $\mathbf{d}_s$ has to be steeply increased with low values of $[\sigma_s]_{\text{lowest}}$. Consequently, the variations of $\mathbf{d}_s$ are defined as

$$\mathbf{d}_s = \mathbf{d}_s^{\min} + c \log\left(1 + [\sigma_s]_{\text{lowest}}^2\right). \tag{2}$$

$c = 6$ was found to be a good compromise for processing our data. To ensure the regularity of the neighboring size image, a 3×3 Gaussian filter is applied over the $\mathbf{d}_s$ image.

III. Predictive Filter for Estimating a DTM

A. General Framework

In its traditional formulation, a predictive filter relates a state prediction of the system at a certain step given the previous state with an effective measurement of the state vector via their respective uncertainties. The final state value is a linear combination of both the prediction and the measurement [15]. Our system is designed to propagate following a spatial scale (cf., Section II-B). It is represented at a site k as an $n \times 1$ state vector, $n \in \mathbb{N}$, $\mathbf{x}_k = (x_{1,k}, \ldots, x_{n,k})^T$, and an $n \times n$ covariance matrix $\mathbf{S}_k$. It evolves depending on a state relation and a measurement relation.

The state relation considers the transition from a site k given its local environment $\mathbb{V}_k$ (Fig. 3) to the following site $k+1$ and is represented as an $n \times m$ transition matrix denoted $\mathbf{\Phi}_{\mathbb{V}_k}$. If $\preccurlyeq$ is an ordered relationship along the trajectory of our system, $\mathbb{V}_k = \{s_i / s_i \preccurlyeq k\}_i$ is defined as the set of neighboring sites s_i (eight-connectivity topology) of $k+1$ which have already been visited (see Fig. 3). We have card $(\mathbb{V}_k) \leq 8$ and $k+1 \notin \mathbb{V}_k$.

Equation (3) (state relation) describes a model of the terrain topography generation seen as a recursive process. Here, we consider that the topography (terrain value and local slope) value at site $k+1$ has a strong correlation with the neighboring sites of $\mathbb{V}_k$. This correlation is translated through the transition matrix $\mathbf{\Phi}_{\mathbb{V}_k}$ as a mean operator on $\mathbb{V}_k$. The state relation can be written as

$$\mathbf{x}_{k+1|\mathbb{V}_k} = \Phi_{\mathbb{V}_k} \mathbf{x}_{\mathbb{V}_k} + \mathbf{w}_k \tag{3}$$

Require: $(\mathbf{x}_{1|0}, \mathbf{S}_{1|0})$ ▷ Initialization
for all site k **do**
  ▷ MEASUREMENT
  $(\mathbf{y}_k, \mathbf{R}_k)$
  ▷ CALCULATION OF THE GAIN
  $\mathbf{K}_k = \mathbf{S}_{k|\mathbb{V}_{k-1}}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{S}_{k|\mathbb{V}_{k-1}}\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$
  ▷ CORRECTION
  $\mathbf{x}_k = \mathbf{x}_{k|\mathbb{V}_{k-1}} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|\mathbb{V}_{k-1}})$
  $\mathbf{S}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{S}_{k|\mathbb{V}_{k-1}}$
  ▷ PREDICTION
  $\mathbf{x}_{k+1|\mathbb{V}_k} = \mathbf{\Phi}_{\mathbb{V}_k}\mathbf{x}_{\mathbb{V}_k} + \mathbf{w}_k$
  $\mathbf{S}_{k+1|\mathbb{V}_k} = \mathbf{\Phi}_{\mathbb{V}_k}\mathbf{S}_k\mathbf{\Phi}_{\mathbb{V}_k}^T + \mathbf{Q}_k$
end for

Fig. 4. Recursive implementation of a predictive filter for DTM generation.

where $\mathbf{w}_k$ is a white noise with covariance $\mathbf{Q}_k$ and $\mathbf{x}_{\mathbb{V}_k}$ is a $\max(\text{card}(\mathbb{V}_k), 8) * \dim(\mathbf{x}_k)$ vector with

$$\mathbf{x}_{\mathbb{V}_k} = \left(\underbrace{x_{1,s_1}, \ldots, x_{1,s_8}}_{X_1}, \ldots, \underbrace{x_{n,s_1}, \ldots, x_{n,s_8}}_{X_n}\right)^T. \quad (4)$$

Let us define the line 1×8 matrix $(\mathbb{1}(X))_{\mathbb{V}_k}$, where $X \in \{X_1, \ldots, X_n\}$ (4), as

$$(\mathbb{1}(X))_{\mathbb{V}_k} = \frac{1}{\text{card}(\mathbb{V}_k)}\left(\mathbb{1}_{\mathbb{V}_k}(s_1), \ldots, \mathbb{1}_{\mathbb{V}_k}(s_8)\right)_X \quad (5)$$

where the indicative function $\mathbb{1}_{\mathbb{V}_k}(s_i)$ is defined as

$$\mathbb{1}_{\mathbb{V}_k}(s_i) = \begin{cases} 1, & \text{if } s_i \in \mathbb{V}_k \\ 0, & \text{if not.} \end{cases} \quad (6)$$

We write the $n \times \dim(\mathbf{x}_{\mathbb{V}_k})$ transition matrix $\mathbf{\Phi}_{\mathbb{V}_k}$ as a block diagonal matrix

$$\Phi_{\mathbb{V}_k} = \begin{pmatrix} (\mathbb{1}(X))_{\mathbb{V}_k} & 0 & \cdots \\ 0 & \ddots & 0 \\ \vdots & 0 & (\mathbb{1}(X_n))_{\mathbb{V}_k} \end{pmatrix}. \quad (7)$$

The measurement equation of our system is written as follows:

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \quad (8)$$

where $\mathbf{y}_k$ is the measurement vector of dimension n, $\mathbf{H}_k$ is a state-measurement matrix (identity matrix here), and $\mathbf{v}_k$ is a measurement noise (white noise) with a covariance matrix $\mathbf{R}_k$.

A predictive filter consists in the following four steps [15]: 1) the measurement of $\mathbf{y}_k$ and its uncertainty $\mathbf{R}_k$; 2) the calculation of the gain $\mathbf{K}_k$ (also called the Kalman gain); 3) the effective value assignment to the state vector $\mathbf{x}_k$ as a linear combination of a previous estimate of the state value $\mathbf{x}_{k|\mathbb{V}_{k-1}}$ and a function of the measurement $\mathbf{y}_k$; and finally 4) a predicted state value $\mathbf{x}_{k+1|\mathbb{V}_k}$ to the next site $k+1$. Fig. 4 shows a recursive implementation of a predictive filter.

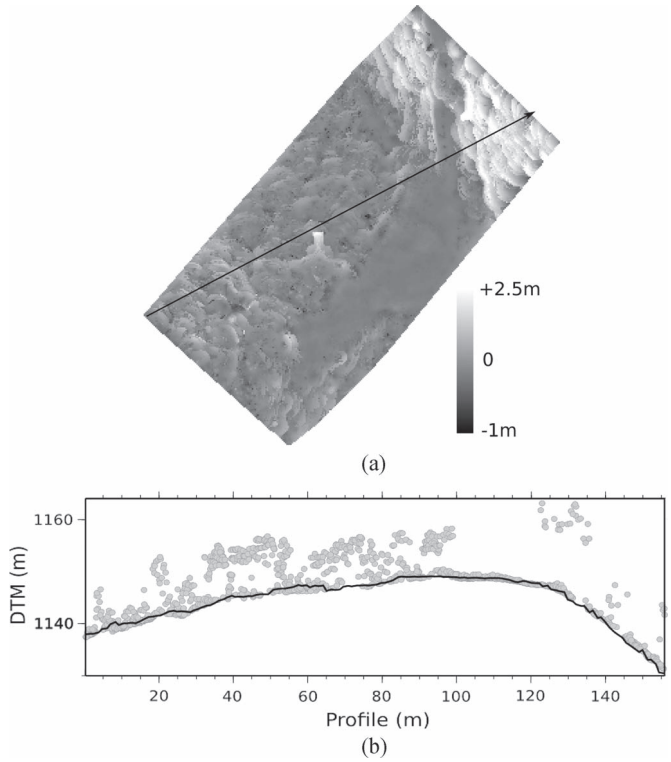


Fig. 5. Influence of the slope on the DTM estimate. (a) Altimetric difference between two DTMs computed with and without the integration of the local slope. Discrepancies are coded in a gray level scale from -1 to 2.5 m. (b) Profile along the black arrow of the DTM calculated with the local slope. The gray points are the lidar points.

B. Estimating the Local Slope

The terrain height estimation depends both on the local slope and on the local neighborhood $(\mathcal{V}_s)$ extension of a site s. Fig. 5 (top) shows the difference between two DTMs calculated with and without the integration of the local slope and a profile (bottom) of a DTM integrating the computation of the local slope. We can notice on this example that the elevation differences can reach 2.5 m when the slope is not taken into account. Additional results are discussed in Section V-C.

The quality of the local plane estimation depends on the lidar point geostatistics within $\mathcal{V}_s$ and defines the terrain height relevancy, as well as its uncertainty. We estimate a plane $n_x x + n_y y + n_z z + d = 0$ with $(n_x, n_y, n_z) \in [-1, 1]$ and $d \in \mathbb{R}$ from lidar points belonging to the first mode of the elevation distribution of $\mathcal{V}_s$. The first mode contains the set of lidar points of the lowest altitude that may belong to the terrain. It is the entire distribution if only one mode has been detected. The 3-D plane estimation is then performed in the barycentric frame. A robust M-estimator has been used with an L_p norm [16]. This algorithm is implemented as an iterative reweighted least square system. The calculation of such a local tangent plane considering each site as independent of its local environment has shown incoherences between neighboring planes. Fig. 6 shows the (n_x, n_y, n_z) components of the normal vectors calculated independently on each site. These components are noisy and make incoherences appear on the terrain topography. These side effects can be overcome using the predictive filter to estimate the local slope that takes into account

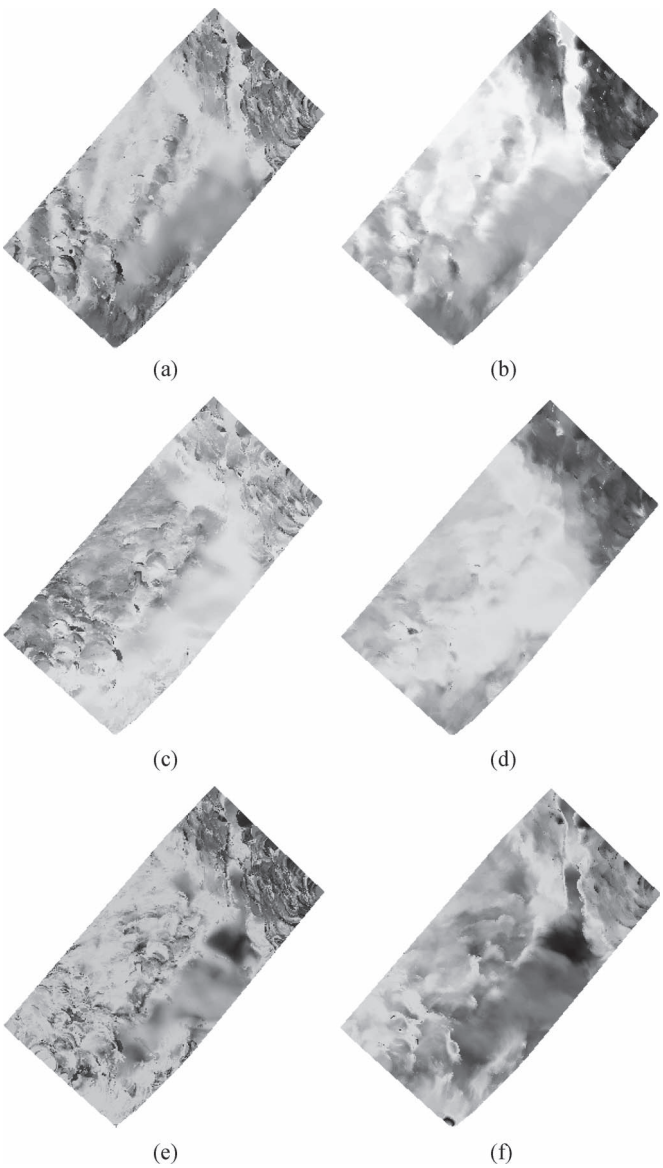


Fig. 6. Local normal vectors calculated on a 1-m resolution grid and displayed in gray level scale. The dynamic of the images has been modified for display purposes. [(a), (c), and (e)] Correspond to the three components calculated independently on each site. [(b), (d), and (f)] Correspond to the three components calculated using a predictive filter. (a) n_x without filtering. (b) n_x with filtering. (c) n_y without filtering. (d) n_y with filtering. (e) n_z without filtering. (f) n_z with filtering.

the local environment of each site. Using the notations defined in Section III-A, $\mathbf{y}_k$ and $\mathbf{x}_k$ are the vectors of dimension 4, $\mathbf{x}_{\mathbb{V}_k}$ is a 32-component vector in the case of a 3×3 neighborhood, and $\mathbf{\Phi}_{\mathbb{V}_k}$ is a 4×32 matrix. The measurement $\mathbf{y}_k$ is computed from the M estimator. Here, $\mathbf{v}_k$ is a Gaussian white noise of zero mean and $\sigma^2 = 0.005$. The covariance matrix $\mathbf{R}_k$ of the slope "measurement" is a diagonal matrix corresponding to the 99% confidence interval around each parameter estimate.[1]

[1]For a linear system $AX = B$, where A is the design matrix, B is the data, and X is the parameter vector, the covariance matrix of the estimate $\hat{X}$ can be written as $C = \sigma^2(A^T A)^{-1}$. The $p\%$ confidence interval is written for each component of $\hat{X}$ as $x_j = \hat{x}_j \pm t(n-m, p\%)\sqrt{C_{jj}}$, where $t(n-m, p\%)$ is the value of the Student coefficient with $n-m$ degrees of freedom at a level of confidence p.

The prediction relation can be rewritten as

$$\mathrm{x}_{k+1|\mathbb{V}_k} = \frac{1}{\mathrm{card}(\mathbb{V}_k)} \sum_{s_i \in \mathbb{V}_k} \begin{pmatrix} n_x(s_i) \\ n_y(s_i) \\ n_z(s_i) \\ d(s_i) \end{pmatrix} \mathbb{1}_{\mathbb{V}_k}(s_1) + \mathbf{w}_k. \quad (9)$$

The variance prediction on the slope $\mathbf{S}_{k+1|\mathbb{V}_k}$ is set to the mean value of the precalculated neighboring slope components. Fig. 6 shows the local normal vectors calculated on a 1-m resolution grid and displayed in gray level scale. The dynamic of the images has been modified for display purposes. They show the improvement of the normal vector calculation in all three coordinates in terms of regularity using a predictive filter with regard to a calculation performed independently on each site.

C. Estimating the Terrain Elevation

Again, the first mode of the elevation distribution is extracted. We then project the lidar points belonging to the first mode into the local frame centered on site k and oriented following the local normal vector $(n_x, n_y, n_z)^T$ calculated previously. For a site k, we denote $\mathcal{F}_k$ the lidar point set belonging to the first mode and projected into the local frame. The measurement $\mathbf{y}_k$ is set to the elevation weighted (planimetric Euclidean distance) mean of the points belonging to $\mathcal{F}_k$.

The uncertainty associated to the terrain height depends on the lidar point statistics of the first mode in the local frame $(n_x, n_y, n_z)^T$. We consider the terrain estimate accurate at one standard deviation of the elevation distribution of $\mathcal{F}_k$. The predictive filter then manages the uncertainty propagation. Here, $\mathbf{v}_k$ is a Gaussian white noise of zero mean and $\sigma^2 = 0.01$, accuracy of the lidar system. Considering ξ_i as the inverse of the planimetric Euclidean distance between a site k and a lidar point $l_i \in \mathcal{F}_k$, we can write the following state equations for this system:

$$\mathbf{y}_k = \frac{1}{\sum_j \xi_j} \sum_i \xi_i \, z_{l_i} + \mathbf{v}_k \quad (10)$$

$$\mathbf{R}_k = \underset{\forall l_i \in \mathcal{F}_k}{\mathrm{variance}} \left(z_{l_i} \right). \quad (11)$$

The terrain value prediction $\mathbf{x}_{k+1|\mathbb{V}_k}$ consists of the mean of the estimated heights of the neighboring sites, which are already processed. The variance prediction on the terrain height $\mathbf{S}_{k+1|\mathbb{V}_k}$ is set to the mean value of the neighboring terrain elevations, which are already processed.

The predictive filter presented previously tends to smooth the terrain surface. Nevertheless, the lidar points describe the topography with much more details, those interesting for a fine description of microrelief. The issue becomes a fusion problem, where a coarse DTM has to incorporate the 3-D point information to become more precise but still regular. This problem can be modeled in a Bayesian framework [17].

IV. Markovian Regularization

A. Background

In a probabilistic framework, an image is a set S of sites s, where the gray level is a descriptor of each site. An image is considered as a realization of a random field $X = (X_s)_{s\in\mathbb{N}^2}$, where X_s is a random variable of the values in $E^{\mathrm{card}(S)}$ [18]. A random field is, therefore, a measurable mapping $X : \Omega \longrightarrow E^{\mathrm{card}(S)}$ associated to a complete measurable space $(\Omega, \mathcal{F}, \mathbf{P})$. This model is described by the probability law $\mathbf{P}(X = x)$, the event x to be a realization of X. As usual in image processing, we will consider the particular Markov random fields (MRFs). In an MRF, the value of a site only depends on its local environment through a neighboring system $\mathbb{V}$ defined as [19]

$$\begin{cases} s \notin \mathbb{V}(s) \\ \forall r \in S\backslash\{s\}, \quad s \in \mathbb{V}(r) \Leftrightarrow r \in \mathbb{V}(s). \end{cases} \tag{12}$$

In the case of an MRF, we have

$$\begin{aligned} \forall x \in \Omega \quad \forall s \in S, \quad & \mathbf{P}\left(X_s = x_s | X_r = x_r, r \in S \setminus \{s\}\right) \\ & = \mathbf{P}(X_s = x_s | X_r = x_r, r \in \mathbb{V}_s). \end{aligned} \tag{13}$$

The Hammersley–Clifford theorem argues the equivalence between an MRF and a Gibbs field. The *a priori* probability of a random variable X can therefore be explicitly written as

$$\mathbf{P}(X = x) = \frac{1}{Z} e^{-\mathcal{U}(x)} \tag{14}$$

where $Z = \sum_{x\in\Omega} e^{-\mathcal{U}(x)}$ is a normalization constant and $\mathcal{U}$ is an energy function (which has the properties to be decomposed into local energies) defined as

$$\begin{aligned} \mathcal{U} : \Omega &\longrightarrow \mathbb{R} \\ x &\longrightarrow \mathcal{U}(x) = \sum_{c\in\mathcal{C}} \mathcal{U}_c(x). \end{aligned} \tag{15}$$

We propose a Bayesian model for regularizing the terrain surface in order to introduce *a priori* knowledge on the model. The Bayes' law, which relates *a priori* and conditional probability, is defined as

$$\mathbf{P}(X|D) = \frac{\mathbf{P}(D|X)\mathbf{P}(X)}{\mathbf{P}(D)} \propto \mathbf{P}(D|X)\mathbf{P}(X). \tag{16}$$

The Bayesian model is related to the inverse problem of how to retrieve the best configuration $\hat{x}$ knowing observations D. We look for the maximum *a posteriori* (MAP estimation) defined as

$$\hat{x}_{\mathrm{MAP}} = \arg\max_{x\in\Omega} \mathbf{P}(X = x|D) \tag{17}$$

that can be written as

$$\hat{x}_{\mathrm{MAP}} = \arg\min_{x\in\Omega} \left(-\log\left(\mathbf{P}(D|X = x)\right) - \log\left(\mathbf{P}(X = x)\right)\right). \tag{18}$$

Under the Markovian hypothesis, solving (18) is equivalent to globally minimizing an energy $\mathcal{E}$, sum of a data term $\mathcal{E}_d$ and a regularization term $\mathcal{E}_r$ (Fig. 7); $\lambda \in \mathbb{R}$ is a weighting parameter

$$\hat{x}_{\mathrm{MAP}} = \arg\min_{x\in\Omega} (\underbrace{\mathcal{E}_d + \lambda\mathcal{E}_r}_{\mathcal{E}}). \tag{19}$$

In our problem, $\hat{x}_{\mathrm{MAP}}$ represents the optimal altitude of a site s.

while $\left|\frac{\mathcal{E}^{\mathrm{it}+1} - \mathcal{E}^{\mathrm{it}}}{\mathcal{E}_{\min}}\right| > \epsilon$ **do**
  for all site s **do**
    while $\mathcal{E}_s(x_s^{t+1}) < \mathcal{E}_s(x_s^t)$ **do**
      $x_s^{t+1} = x_s^t + \delta\,\mathrm{sgn}\left(\mathcal{E}_s(x_s^t - \delta) - \mathcal{E}_s(x_s^t + \delta)\right)$
    end while
  end for
end while

Fig. 7. Details of the optimization step of the Markovian regularization.

B. Definition of the Energy

1) Data Term: The data term at site s depends on the Euclidean distance between the terrain height x_s and a value ζ_s considered as a surface attractor. We consider that lidar points located within a buffer zone over the initial DTM may contribute to a better description of the reality of the terrain surface. However, the predictive filter described previously generates not only a terrain surface but also a standard deviation σ_{x_s} associated to these measurements. Therefore, this buffer zone is defined for each site as proportional to the associated standard deviation. Based on the lidar point set $\mathcal{N}_s$ corresponding to a site s and within the tolerance layer at site s, ζ_s is finally calculated as the average altitude of lidar points belonging to $\mathcal{N}_s$. Moreover, when no lidar points are detected within the tolerance layer, the Kalman surface is used. The Euclidean distance is then weighted by the number of lidar points $(\mathrm{card}(\mathcal{N}_s))$ if there are any. If not, the weight ω is set to one.

We define the set $D = \{s/ - q\sigma_{x_s} \leq x_s - \zeta_s \leq q\sigma_{x_s},\, q \in \mathbb{N}\}$. The Euclidean distance has to be minimal so that the final DTM should be as near as possible to the laser measurements considered as the ground. We therefore define $\mathcal{E}_d$ as

$$\begin{aligned} \mathcal{E}_d(D|X = x_s) &= \omega(\zeta_s - x_s)^2 \\ &\text{with } \omega = \max\left(1, \mathrm{card}(\mathcal{N}_s/s \in D)\right). \end{aligned} \tag{20}$$

2) Regularization Term: The regularization term aims to compensate the effect of the data term so that the final surface should not be too noisy. This term depends on the intrinsic geometry of the surface [20].

Let h be the surface defined as

$$\begin{aligned} h : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longrightarrow (x, y, z) = h(x, y). \end{aligned}$$

A second-order Taylor expansion at point $\mathbf{u}_0 = (x_0, y_0)$ with $\mathbf{u} = [x, y]^T$ can be written

$$\begin{aligned} h(\mathbf{u}) = & \underbrace{h(\mathbf{u_0}) + (\mathbf{u} - \mathbf{u_0}).\nabla_{\mathbf{u_0}} h}_{\Pi_0(\mathbf{u})} \\ & + \frac{1}{2}(\mathbf{u} - \mathbf{u_0})\mathbf{H}(\mathbf{u} - \mathbf{u_0})^T + o\left(\|(\mathbf{u} - \mathbf{u_0})\|^2\right) \end{aligned} \tag{21}$$

TABLE I
REVIEW OF THE PARAMETERS

Process	Parameter	Value
Variance	σ_{min}	1m
Adapting local neighboring system	c	6
Vertical distribution	class width	0.3 m
Predictive Kalman	-	-
Markovian regularization	q (Buffer Zone) λ (regularization parameter) δ (steepest gradient step)	6 0.1 m 0.01 m

where Π_0 is the tangent plane to h in $\mathbf{u}_0$ and $\mathbf{H} = \begin{pmatrix} \frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 h}{\partial x \partial y} \\ \frac{\partial^2 h}{\partial x \partial y} & \frac{\partial^2 h}{\partial y^2} \end{pmatrix}$ is the Hessian matrix of h. This matrix describes the local properties of the surface curvature.

We define the regularization term as a function of the trace and the determinant of the Hessian matrix. The trace describes the local convexity of the surface, while the determinant is linked to the shape of the surface with regard to its tangent plane (parabolic, elliptic, and hyperbolic). We therefore define $\mathcal{E}_r$ as

$$\mathcal{E}_r = \alpha_1 \mathrm{tr}(H)^2 - \alpha_2 \det(\mathbf{H})\alpha_1, \qquad \alpha_2 \in \mathbb{R}^+ \tag{22}$$

$$= \frac{\alpha_2}{2}\left(\left(\frac{\partial^2 h}{\partial x^2}\right)^2 + \left(\frac{\partial^2 h}{\partial y^2}\right)^2\right) + \alpha_2\left(\frac{\partial^2 h}{\partial x \partial y}\right)^2 + \left(\alpha_1 - \frac{\alpha_2}{2}\right)\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right)^2. \tag{23}$$

This energy is designed so that its convexity should be managed for optimization purposes. Indeed, in the case of estimating a fine terrain surface, there are not any forbidden natural shapes. A repulsive term is, therefore, not appropriate. If the constraints

$$\alpha_2 \geq 0 \qquad \alpha_1 \geq \frac{\alpha_2}{2} \tag{24}$$

are applied, the energy becomes convex. In practice, $\alpha_1 = 1$, and $\alpha_2 = 1/2$.

3) Optimization: Since $\mathcal{E}$ is a separable energy in a Markovian formulation of the problem, the optimization is based on an iterated conditional mode algorithm [21]. It is split into subproblems, each of them applied to a particular DTM site (15). The related local energy at site s is denoted $\mathcal{E}_s$. An iteration toward the global minimum of $\mathcal{E}$ is denoted $\mathcal{E}^{\mathrm{it}}$ in the following algorithm. Because of the convexity of $\mathcal{E}$, the minimum is locally found using a steepest gradient method with constant step δ. The following algorithm summarizes the optimization steps, where index t refers to an iteration on a particular site toward the local minimum. The convergence is obtained when the global energy has no significant changes within ϵ.

V. RESULTS AND DISCUSSION

A review of the parameters is presented in Table I.

Fig. 8. Orthorectified image of the area whereon the results are presented.

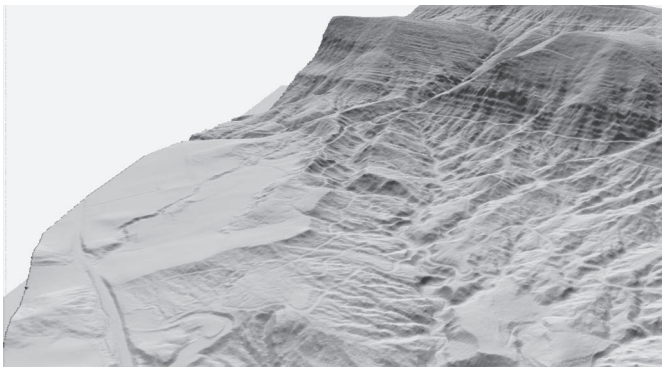


Fig. 9. Three-dimensional bird's eye view of a shaded DTM over Le Brusquet, France.

A. Data Set

The algorithm has been tested on a large data set over Le Brusquet in France. It is a mountainous area with steep relief and vegetation. The whole data consist of 23×10^6 laser points.

The data acquisition was performed in April 2007 using a RIEGL LMS-Q560 system. The lidar system operated at a pulse rate of 111 kHz. The flight height was around 500 m, leading to a footprint size of about 0.25 m. The point density was about 5 pts/m^2.

As mentioned in Section II-A, the entire survey is paved into lidar point subsets of roughly 1.5×10^6 points (Fig. 2) with 40-m overlap. The DTM is processed independently on each subset. The entire final DTM is finally reconstructed by mosaicking the different DTMs.

B. Results

We present the results obtained from a 2.6 km $\times$ 1.6 km area (2.3×10^6 points) subset of Le Brusquet's data (Fig. 8). Fig. 9 shows the final DTM over the entire area of Le Brusquet using a 3-D shaded bird's eye view. Small details like roads or microrelief are clearly visible. The final DTM resolution is 1 m. Fig. 10 shows the textured DTM with the corresponding orthorectified aerial image. The mountain flanks are dark due to the sun shadow. Nevertheless, the roads are particularly well reconstructed at the forefront.

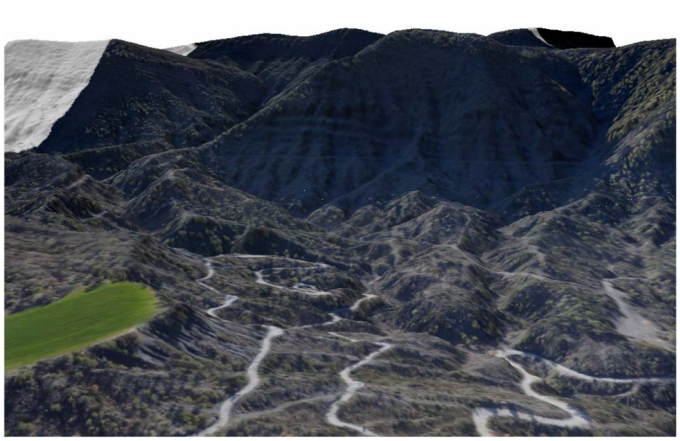


Fig. 10. Three-dimensional bird's eye view of the textured DTM over Le Brusquet, France. The textured image is the orthorectified image.

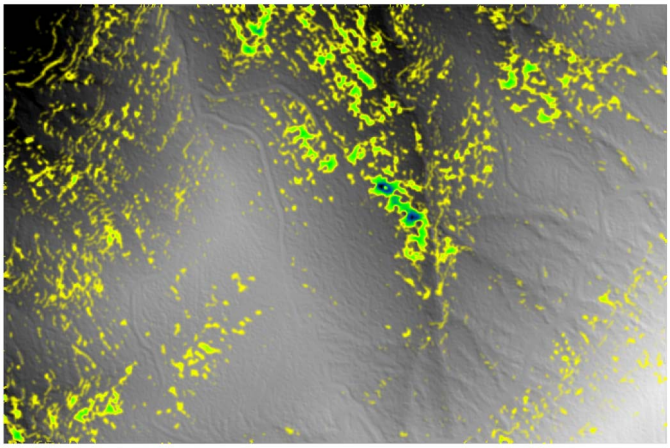


Fig. 11. Superimposition of the $\mathbf{d}_s$ values and of the final DTM. $\mathbf{d}_s$ are presented in a color scale from (transparent and yellow) low values to (green) medium and (dark blue) highest values.

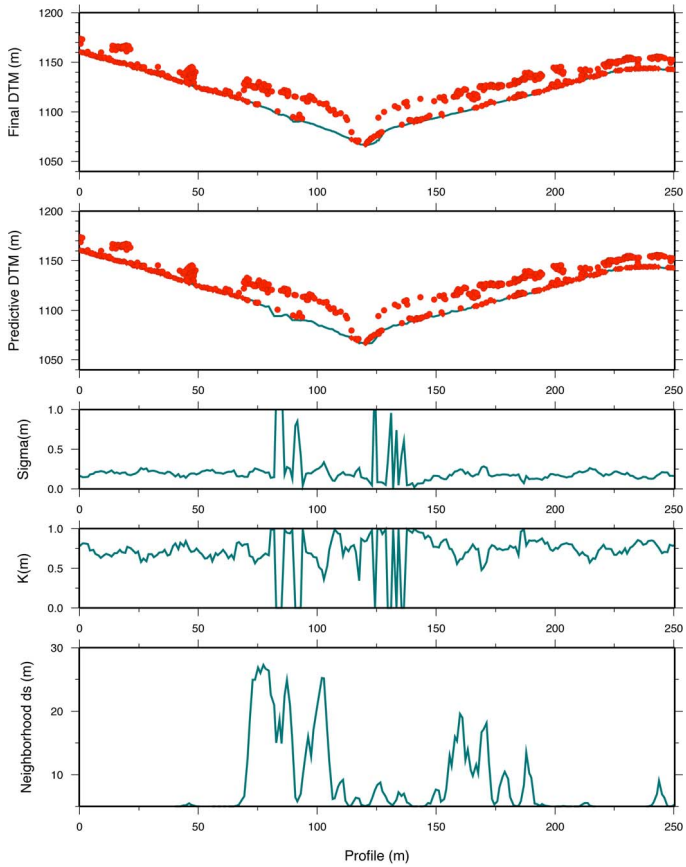


Fig. 12. Profiles of the different outputs of the algorithm. (From bottom to top) Diameter of the cylinder used to estimate the terrain altitude in the Kalman filter, Kalman gain, standard deviation of the terrain height estimate, profile of the Kalman DTM, and, finally, of the regularized DTM.

C. Discussion

1) On the Predictive Filter: Fig. 11 shows the values of $\mathbf{d}_s$ superimposed on the final DTM. Figs. 12 and 13 show two profiles of different outputs of the algorithm following a coherent x-axis. Regarding the local neighborhood diameter $\mathbf{d}_s$, one can remark on both profiles that the diameter increases while lidar ground points are missing and is kept minimal while not. We decided to manage *a priori* information beforehand by generating a landscape predictor mask and to adapt the local environment depending on this predictor. If other predictors would be required (urban areas, watered areas, etc.), we would need to relate the geostatistics of lidar points to these predictors. The behavior of the statistics of the lidar data depending on the landscape has still to be investigated in terms of reflectivity and geometry.

Generally, in a natural landscape and at a global scale, the probability of lidar pulses to be transmitted through the canopy cover is higher than their probability to be reflected by the canopy top or by the lower canopy level. Therefore, the first elevation mode is likely to belong to the terrain, and $\mathbf{d}_s$ is kept minimum. Thus, the terrain measurement and the prediction are almost similar at a global scale, which is well adapted for linearly combining both values. It is confirmed by the mean value of the Kalman gain along the profile in Fig. 13, where $\overline{\mathbf{K}} = 0.5$. When $\mathbf{d}_s$ increases, i.e., when the probability of finding ground points increases, the terrain measurement is less accurate in the case of relief, and the surface is smoothed. The prediction is therefore privileged. Moreover, it is clear that the adapting neighborhood based on variance measurements fails when encountering steep topographic relief like cliffs. Nevertheless, breaklines can be inserted to serve as attractors in the regularization step.

The uncertainty associated to the DTM follows the opposite variations of the gain's ones. A high gain weighs more the measured values with regard to the predicted one and enhances the DTM accuracy. This uncertainty varies around the variance of the white noise considered as the accuracy of the lidar measurement itself (Section III-C). However, it is certainly underestimated in terms of absolute accuracy. It should be improved with prior information. Using a multiple class landscape predictor/descriptor calculated from alternative data sources would improve our knowledge of the expected uncertainty. For instance, the density and the nature of the vegetation influence the penetration rate of a lidar beam within the canopy.

2) On the Markovian Regularization: The results of the regularization step are presented in the top profiles in Figs. 12 and 13. The improvement in terms of regularity is clearly visible. Fig. 14 shows the image of the differences between the Kalman DTM and the regularized one. One can remark

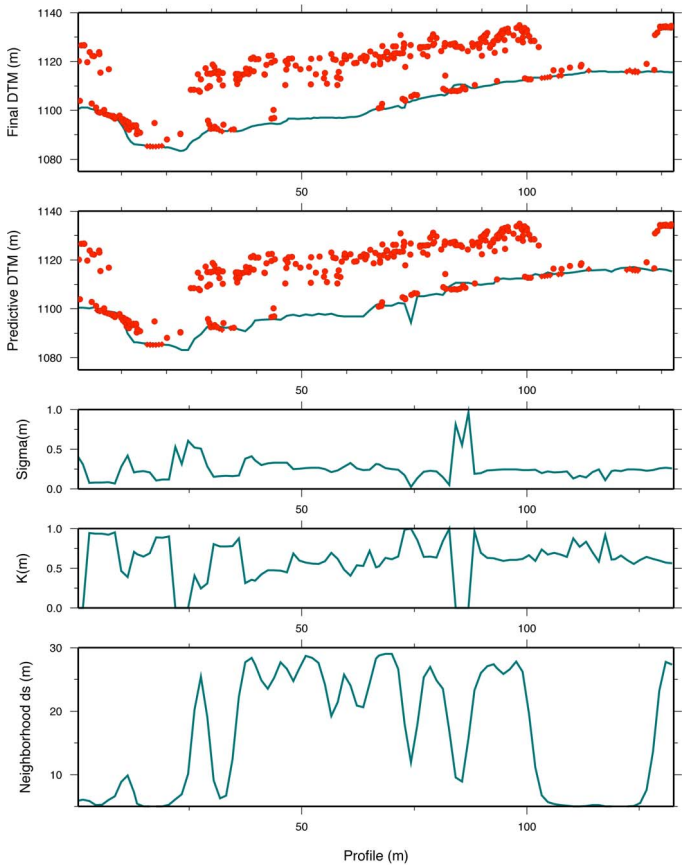


Fig. 13. Profiles of the different outputs of the algorithm. (From bottom to top) Diameter of the cylinder used to estimate the terrain altitude in the Kalman filter, Kalman gain, standard deviation of the terrain height estimate, profile of the Kalman DTM, and, finally, of the regularized DTM.

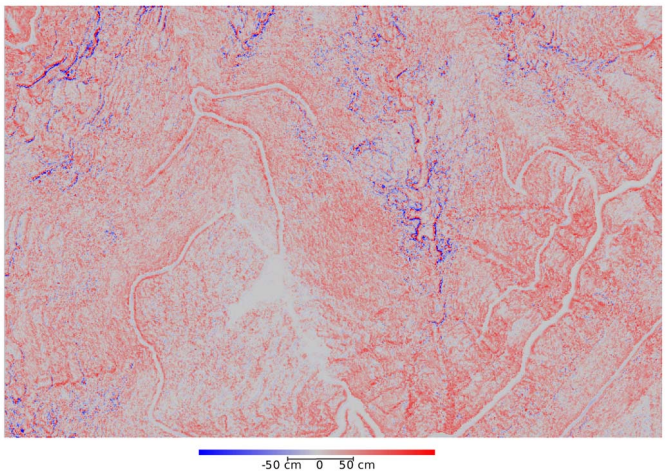


Fig. 14. Image of the differences between the Kalman DTM and the regularized one. The color scale is set so that the red color represents the positive displacements and the blue colors the negative displacements.

that the surface barely evolves when lidar points are considered as attractors (Fig. 15): The Kalman surface is well estimated when there are lidar ground points. Therefore, the data term over such sites is weighted in the energy minimization process by the number of lidar points. Fig. 14 also shows that most of the deformations are positive (red pixels), i.e., that the regularized surface is higher than the Kalman surface. Indeed,

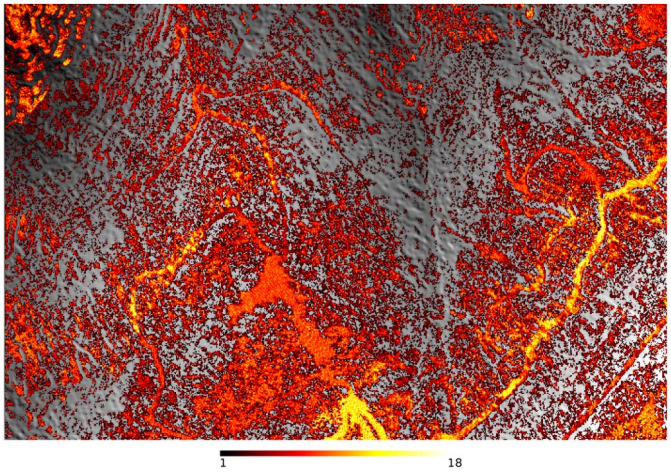


Fig. 15. Image of the number of attractors presented in a red hot color scale and superimposed to a shaded DTM.

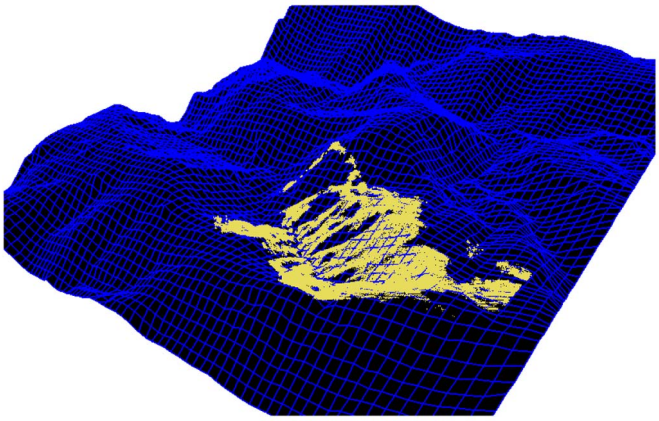


Fig. 16. (Yellow points) Terrestrial lidar points superimposed on the meshed DTM.

the initial surface is meant to reproduce low frequencies of the terrain: In mountainous landscapes, crests and flanks can be underestimated. Moreover, in the case of dense vegetation, the algorithm increases $\mathbf{d}_s$, entailing an estimation of the terrain over a large area. As a result, the terrain is locally smoothed. If lidar attractors are located within the buffer zone (defined as proportional to the accuracy of the Kalman DTM), the surface will be strongly attracted, particularly along breaklines (crests, thalwegs, and path sides). Otherwise, within regions where there are no lidar points (transparent zones in Fig. 15), the surface will evolve mainly following the regularization term of the energy with the Kalman estimation as attractors. In a future work, we would like to integrate other *a priori* information like breaklines in the minimization process.

The algorithm is controlled by two parameters: the weighting parameter of the regularization term λ and the step δ of the steepest gradient algorithm. No significant changes have been noticed when parameter λ belongs to [0, 1]. It is not a critical parameter in the algorithm because the data term is dominant most of the time. δ defines the final accuracy of the DTM. It is set to $\delta = 0.01$ m. As mentioned in Section V-C3, there is no need to take a value that is smaller than the intrinsic accuracy of the lidar points.

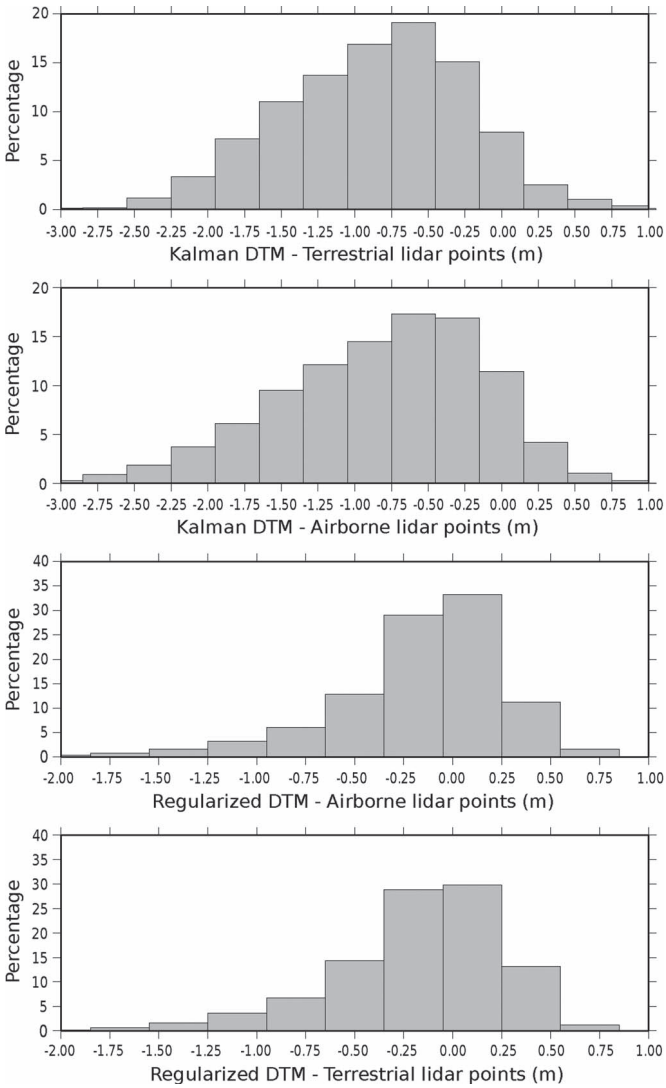


Fig. 17. Distributions of the elevation differences between (from top to bottom) the Kalman DTM and terrestrial lidar, the Kalman DTM and airborne lidar, the regularized DTM and airborne lidar, and the regularized DTM and terrestrial lidar.

3) Comparison With External Data: In order to quantify the accuracy of the final DTM, external data have been acquired over a small area of some hundreds of square meters using a terrestrial lidar (2.8×10^6 points). This area is vegetation free and represents a small valley. Such systems provide a centimetric accuracy along the lidar direction [22]. Fig. 16 shows these terrestrial lidar points superimposed on the meshed DTM. We compared the distribution of the elevation differences between the terrestrial/airborne lidar points and the Kalman/regularized DTMs. The results are presented in Fig. 17 and in Table II. One can notice that the Kalman DTM is, *on average*, lower than both the airborne lidar points and the terrestrial lidar points (-0.81 ± 0.7 m and -0.86 ± 0.63 m, respectively). As already mentioned, the Kalman filter is designed to be robust against outliers (estimation of the terrain over a large neighborhood in the case of high elevation variance) and, therefore, may be underestimated to keep the global shape of the terrain surface.

The distribution of the elevation differences between the terrestrial/airborne lidar points and the regularized DTMs fully justifies the regularization step: The average differences are -0.16 ± 0.43 m and -0.16 ± 0.44 m, respectively. The intrinsic coherence of the terrestrial/airborne lidar points in Table II represents the average distance of the lidar points with regard to a robust local plane. The intrinsic coherence can be observed in Fig. 18, where a "noise" clearly appears on the airborne lidar points. After close investigations of the airborne point clouds, it appears that a misadjustment between two overlapping strips explains the higher intrinsic coherence of the airborne lidar data. Nevertheless, the final DTM leads to the same altimeter difference between airborne and terrestrial lidar which confirms the robustness of the algorithm.

TABLE II
STATISTICS OF THE DTMS WITH REGARD TO TERRESTRIAL AND AIRBORNE LIDAR POINTS

	Aerial Lidar points	Terrestrial Lidar points
Intrinsic coherence	0.14 m	0.03 m
Predictive DTM	-0.81 ± 0.7 m	-0.86 ± 0.63 m
Final DTM	-0.16 ± 0.44 m	-0.16 ± 0.43 m

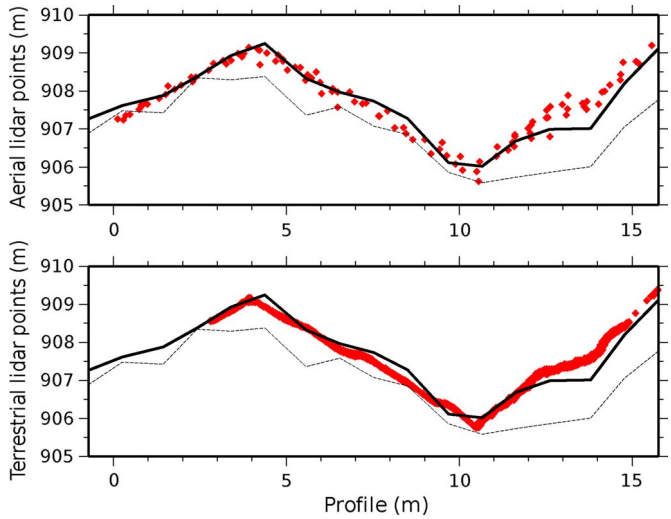


Fig. 18. Profiles of the (red points) terrestrial and airborne lidar data, as well as the (dashed line) Kalman and (black line) regularized DTM.

VI. CONCLUSION

We have presented in this paper a methodology for generating a DTM from 3-D lidar data. It is a full operational system which works on complex and very large data sets. The main contribution of this work consists in using the coherence of the terrain in a predictive filter paradigm to deal with the irregular point cloud representation of the topography. The predictive filter is designed to integrate both the local slope and landscape predictors into the estimation of the ground altitude.

In this paper, results are presented on mountainous landscapes with vegetation, but the methodology can be extended to different landscapes such as rural, urban, and peri-urban areas provided that appropriate landscape predictors are available. Moreover, the improvement of the regularization process imposes the modeling and the integration of breaklines (crests, thalwegs, mountain paths, building footprints, etc.) in the global

energy. Among other applications, the quality of the DTM not only manages the relevancy of the lidar point classification for studying off-ground features but also ensures the quality of a reality-based landscape rendering for GIS-based applications.

ACKNOWLEDGMENT

The authors would like to thank GIS Draix for providing the lidar data and helping in the ground truth surveys; INSU for its support to GIS Draix through the ORE program; A. Chauve and O. Tounaire (MATIS Laboratory, IGN) for the implementation of the image viewer; and M. Bredif (MATIS Laboratory, IGN) for the implementation of the 3-D viewer.

REFERENCES

[1] S. Roy and I. Cox, "A maximum-flow formulation of the n-camera stereo correspondence problem," in *Proc. IEEE Int. Conf. Comput. Vis.*, Bombay, India, Jan. 1998, pp. 492–499.

[2] M. Pierrot-Deseilligny and N. Paparoditis, "A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery," in *Proc. ISPRS Conf. Topographic Mapping From Space (With Special Emphasis on Small Satellites)*, Feb. 2006. [CD-ROM].

[3] M. Pierrot-Deseilligny, MicMac, un logiciel pour la mise en correspondance automatique dans le contexte gographique: Institut Géographique National, 2008. [Online]. Available: http://www.micmac.ign.fr/

[4] E. Ahokas, H. Kaartinen, and J. Hyyppa, "A quality assessment of airborne laser scanner data," in *Proc. ISPRS Workshop III/3 '3D Reconstruction From Airborne Laserscanner and InSAR'—International Archives of Photogrammetry and Remote Sensing*, Dresden, Germany, Oct. 2003, vol. XXXIV, pp. 1–7.

[5] ALS60 Brochure, 2008. displayed on 17/03/2009. [Online]. Available: http://www.leica-geosystems.com/corporate/en/ndef/lgs_57629.htm

[6] Gemini 167, 2008. displayed on 17/03/2009. [Online]. Available: http://www.optech.ca/prodaltm.htm

[7] G. Sithole and G. Vosselman, "Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 59, no. 1/2, pp. 85–101, Aug. 2004.

[8] B. Petzold, P. Reiss, and W. Stossel, "Laser scanning—Surveying and mapping agencies are using a new technique for the derivation of digital terrain models," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 2/3, pp. 95–104, Jul. 1999.

[9] W. Eckstein and O. Munkelt, "Extracting objects from digital terrain models," in *Proc. SPIE—Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*, 1995, vol. 2572, pp. 43–51.

[10] K. Zhang, S.-C. Chen, D. Whitman, M. Shyu, J. Yan, and C. Zhang, "A progressive morphological filter for removing nonground measurements from airborne lidar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 4, pp. 872–882, Apr. 2003.

[11] P. Axelsson, "DEM generation from laser scanner data using adaptive TIN models," in *Proc. Int. Arch. Photogramm. Remote Sens.*, 2000, vol. XXXIII, pp. 110–117.

[12] K. Kraus and N. Pfeifer, "Determination of terrain models in wooded areas with airborne laser scanner data," *ISPRS J. Photogramm. Remote Sens.*, vol. 53, no. 4, pp. 193–203, Aug. 1998.

[13] H. Lee and N. Younan, "DTM extraction of lidar returns via adaptive processing," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2063–2069, Sep. 2003. [CD-ROM].

[14] Terrasolid, Terrascan User's Guide, Jyväskylä, Finland, 2006. [Online]. Available: http://www.terrasolid.fi/en/users guide

[15] G. Evensen, *Data Assimilation. The Ensemble Kalman Filter*. New York: Springer-Verlag, Oct. 2006.

[16] G. Xu and Z. Zhang, *Epipolar Geometry in Stereo, Motion and Object Recognition*. Norwell, MA: Kluwer, 1996.

[17] F. Bretar, "Processing fine digital terrain models by Markovian regularization from 3D airborne lidar data," in *Proc. IEEE Int. Conf. Image Process.*, San Antonio, TX, Sep. 2007, pp. IV-125–IV-128.

[18] P. Pérez, "Markov random fields and images," INRIA, Sophia Antipolis, France, Tech. Rep. 1196, 1998.

[19] S. Li, *Markov Random Field Modeling in Computer Vision*. New York: Springer-Verlag, 1995.

[20] P. Fua, "Model-based optimization: An approach to fast, accurate, and consistent site modeling from imagery," in *RADIUS: Image Understanding for Intelligence Imagery*. San Mateo, CA: Morgan Kaufmann, 1997.

[21] J. Besag, "On statistical analysis of dirty pictures," *J. R. Stat. Soc.*, vol. 48, no. 3, pp. 259–302, 1986.

[22] A. Ruiz, W. Kornus, J. Talaya, and J. Colomer, "Terrain modeling in an extremely steep mountain: A combination of airborne and terrestrial lidar," in *Proc. XXth ISPRS Symp.*, Jul. 2004, pp. 281–284.

Frédéric Bretar (M'07) received the Dipl.Eng. degree from the Ecole Nationale des Sciences Géographiques, Paris, France, and the M.S. degree in remote sensing from the University of Paris 7, in 2000. He defended his Ph.D. degree in image and signal processing at TELECOM ParisTech, France, in 2006.

From 2000 to 2002, he was an Invited Researcher with the Norwegian Defense Research Establishment on the interpretation of satellite SAR images and on InSAR data. From 2006 to 2009, he was the Head of the Lidar Group at the MATIS laboratory (Institut Géographique National). He is currently with the Ecole des Ponts ParisTech, France. He is a Specialist of airborne lidar data processing for quantitative cartography (pattern recognition and 3-D modeling), as well as of image processing and photogrammetry. His areas of interest are the processing of lidar data (3-D point cloud and full waveform), as well as the data fusion framework, with a particular interest in forest areas and geophysics.

Dr. Bretar has chaired the WG-III/2 "Processing of Point Clouds" of the International Society of Photogrammetry and Remote Sensing since 2008.

Nesrine Chehata (M'08) received the Dipl.Eng. degree in physics from the Ecole Nationale Supérieure de Physique de Strasbourg, Strasbourg, France, in 2001 and the M.S. degree in image processing from the University of Strasbourg I, Strasbourg, in 2001. She defended her Ph.D. thesis in computer vision and image processing in 2005 at the University of Paris 5, Paris, France.

She worked in collaboration with the Méthodes d'Analyses pour le Traitement d'Images et la Stéréorestitution (MATIS) Laboratory, Institut Géographique National [(IGN); French cartographic institute], St. Mandé, France, and the Centre National d'Études Spatiales [(CNES), French spatial agency]. Since 2006, she has been an Associate Professor in the EGID Institute (environment and geosciences institute), University of Bordeaux 3, Bordeaux, France. She is also an Associate Researcher with the MATIS Laboratory, IGN. Her research interests are pattern recognition and image processing applied to remote sensing data such as airborne or terrestrial lidar data and VHR optical satellite images for environmental applications.

Dr. Chehata is a member of the IEEE Geoscience and Remote Sensing Society.

5.3.1 Classification de dégâts de tempête sur la forêt des Landes

International Journal of Remote Sensing, 2014
Vol. 35, No. 13, 4758–4777, http://dx.doi.org/10.1080/01431161.2014.930199

Taylor & Francis
Taylor & Francis Group

Object-based change detection in wind storm-damaged forest using high-resolution multispectral images

N. Chehata[a,b]*, C. Orny[a,c], S. Boukir[a], D. Guyon[c], and J.P. Wigneron[c]

[a]*EA 4592 G&E, ENSEGID-IPB, University of Bordeaux, Pessac Cedex, France;* [b]*IRD, LISAH UMR 144, Tunis, Tunisia;* [c]*INRA, UMR 1391 ISPA, Villenave d'Ornon, France*

(*Received 2 April 2013; accepted 4 February 2014*)

Natural disasters are generally brutal and may affect large areas, which then need to be rapidly mapped to assess the impacts of such events on ecosystems and to prevent related risks. Ground investigations may be complex, whereas remote-sensing techniques enable a fast regional-scale assessment of damage and offer a cost-effective option for large and inaccessible areas. Here, an efficient, quasi-automatic object-based method for change mapping using high-spatial-resolution (HR) (5–10 m) satellite imagery is proposed. Our contribution comprises two main novelties with respect to similar works in forestry. First, an automatic feature selection process optimizes the image segmentation and classification steps via an original calibration-like procedure. Second, an automatic bitemporal classification enables the separation of damaged and intact areas thanks to a new descriptor based on the level of fragmentation of the obtained regions. The mean shift algorithm is used in both the segmentation and classification processes. The method was assessed in a maritime pine forest using bitemporal HR Formosat-2 multispectral images acquired pre- and post-Windstorm Klaus, which occurred in January 2009 in southwestern France. The binary overall classification accuracy reached 87.8% and outperformed a pixel-based K-means classification with no feature selection. A thematic analysis of the results highlights the correlation between the ages of trees and their sensitivity to wind.

Downloaded by [Inra] at 00:00 10 July 2014

1. Introduction

In a climate change context, wind storms have become more and more frequent. Windfall damages have to be quickly mapped to prevent fire risks and to update the national forest inventory. Although ground investigations are complex due to the obstacles created by fallen trees, remote-sensing techniques enable rapid monitoring of large and inaccessible areas. This study aims to provide a binary map discriminating damaged and non-damaged areas using high-spatial-resolution (HR) (5–10 m) satellite imagery. Bitemporal 8 m multispectral Formosat-2 images, acquired pre- and post-Windstorm Klaus, which occurred on 24 January 2009 in southwestern France, were used for the map.

In the literature, previous studies in forestry have produced low-resolution maps (i.e. minimal mapping unit $\approx$ 1 ha), using medium- or high-spatial resolution sensors and focused essentially on defoliation (Franklin et al. 1995), clear-cut detection (Fraser, Abuelgasim, and Latifovic 2005; Desclée, Bogaert, and Defourny 2006; Huang et al. 2008; Liu et al. 2008; Conchedda, Durieux, and Mayaux 2008), or deforestation (Hayes and Cohen 2007; McRoberts and Walters 2012; Vibrans et al. 2013). Fewer studies have addressed more abrupt and smaller-scale structural changes, such as windfall damage (Meyer, Geldreich, and Yesou

*Corresponding author. Email: nesrine.chehata@ensegid.fr

© 2014 Taylor & Francis

2001; Liu et al. 2008; Kennedy et al. 2009; Wang and Xu 2010; Jonikavičius and Mozgeris 2013). Moreover, the above works have generally utilized pixel-based approaches based on medium-spatial resolution sensors, such as Landsat TM or MODIS.

Bitemporal change detection methods can be based on either post-classification approaches that compare two classifications obtained separately using remotely sensed data from two dates (McRoberts and Walters 2012; Liu et al. 2008), or the direct classification of two sets of remotely sensed data that have been merged into a single dataset (Hayes and Cohen 2007). Post-classification approaches are robust to radiometric differences between images and provide accurate 'from–to' change information (Im and Jensen 2005) but suffer from segmentation and classification error propagation. Most post-classification approaches are pixel-based. Joint-classification approaches provide more information with which to classify small changes. They can rely on machine-learning algorithms using a training set (Im and Jensen 2005), expert knowledge (Fraser, Abuelgasim, and Latifovic 2005), or, alternatively, thresholding, which involves a parametric statistical test (Desclée, Bogaert, and Defourny 2006). In this work, the proposed method is based on bitemporal joint classification of forest damage (damaged vs. non-damaged).

Storm damage assessment using multispectral images can be separated into two approaches: (1) storm damage modelling based on image descriptors and (2) storm damage mapping using image classification methods. In the first approach, the relationships between image descriptors and forest variables are assessed – for example, the relationship between NDVI variation and the proportion of damaged area per pixel (Wang et al. 2010) or between image texture and the percentage of crown loss (King et al. 2005). However, the correlations obtained are insufficient for accurate damage mapping (Olthof, King, and Lautenschlager 2004; King et al. 2005). The second approach includes time-consuming traditional visual interpretation of imagery (Clandillon, Yesou, and Meyer 2003) and few automated methods. Some authors (Olthof, King, and Lautenschlager 2004; Meyer, Geldreich, and Yesou 2001) have proceeded with a traditional pixel-based supervised classification using differencing images and vegetation indices, whereas Schwarz, Steinmeier, and Waser (2001) used a supervised object-oriented classification considering four spectral bands. The proposed method is based on a binary joint classification of forest storm damage.

Given the increasing spatial resolution of new satellites, various methods have been developed to exploit contextual spatial information. Some approaches use regional per-pixel descriptors before a pixel-wise classification (King et al. 2005; Fraser, Abuelgasim, and Latifovic 2005; Im and Jensen 2005; Franklin et al. 1995). Recently, the use of object-based image analysis (OBIA) has increased in the remote-sensing community (Benz et al. 2004; Blaschke 2010). In forestry, object-based classifications are used for structure mapping (Mallinis et al. 2008; Lamonaca, Corona, and Barbati 2008; Johansen et al. 2007; Mäkelä and Pekkarinen 2001; Pasher and King 2009) and for change detection (Desclée, Bogaert, and Defourny 2006; Schwarz, Steinmeier, and Waser 2001; Conchedda, Durieux, and Mayaux 2008; Wulder et al. 2008; Chehata et al. 2011). The OBIA approach is based on a segmentation process that combines spatial and spectral information to group pixels into homogeneous regions before their classification using new object descriptors. These object descriptors can be geometrical and textural (Fraser, Abuelgasim, and Latifovic 2005) or temporal (Desclée, Bogaert, and Defourny 2006; Chehata et al. 2011). Bitemporal object-based analysis has the advantage of reducing misregistration and shadowing effects compared with a pixel-based approach (Johansen et al. 2010).

Various segmentation algorithms are used in the forestry literature, the most common being a region-growing algorithm that minimizes intra-region variance (Kim, Madden, and Warner 2008; Desclée, Bogaert, and Defourny 2006) generally using Definiens

Downloaded by [Inra] at 00:00 10 July 2014

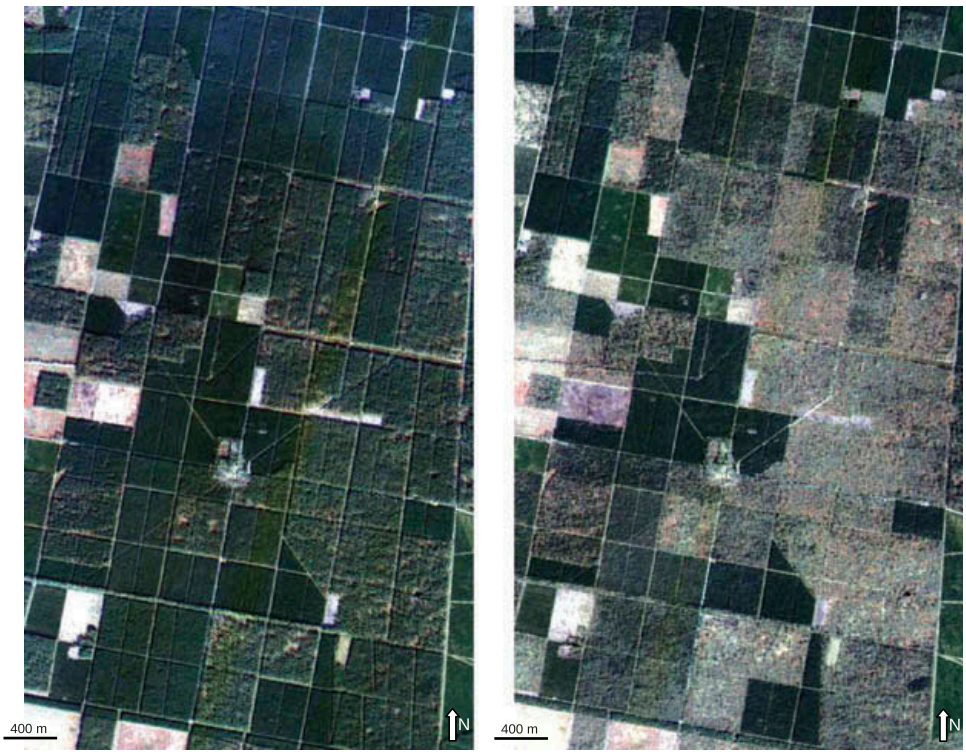
Imaging 2004 software (Benz et al. 2004). Mäkelä and Pekkarinen (2001) used the Narendra and Goldberg segmentation algorithm to group pixels according to a heterogeneity measure based on local gradients. Liu et al. (2008) used a supervised Markov random field for segmentation and classification with good results for detection of areas of severe and large forest change. The mean shift (MS) algorithm has also exhibited satisfactory results for remote-sensing image segmentation applications such as individual tree crown delineation from aerial images (Wang, Soh, and Shultz 2006) or agricultural land delineation (Ozdarici and Akyurek 2012). To our knowledge, the MS algorithm has not been used for forest cover mapping. It will be detailed in Section 3.2.

The proposed method is an object-based, bitemporal classification that maps storm damage at a fine spatial scale. It is nearly automatic, requiring limited training data to rapidly map damage over a wide area.

2. Study area and data

The Nezer forest covers approximately 60 km^2 and is located near the Atlantic coast in southwest France, within a large European maritime pine *(Pinus pinaster* Ait.*)* forest. The forest is made up of rectangular stands of pine trees of similar age and height. Stand age varies from seedling to 55 years, density from 150 to 3000 trees ha^{-1}, and tree height from 0 to 25 m, as reported in Garestier et al. (2009). The maximum crown diameter reaches approximately 8 m for the oldest stands (Guyon et al. 2005).

Bitemporal Formosat-2 images from pre- and post-Windstorm Klaus, acquired on 22 December 2008 and 4 February 2009, respectively, were used (Figure 1). The images



(*a*) (*b*)

Figure 1. Formosat-2 multispectral images acquired before and after Windstorm Klaus.

Downloaded by [Inra] at 00:00 10 July 2014

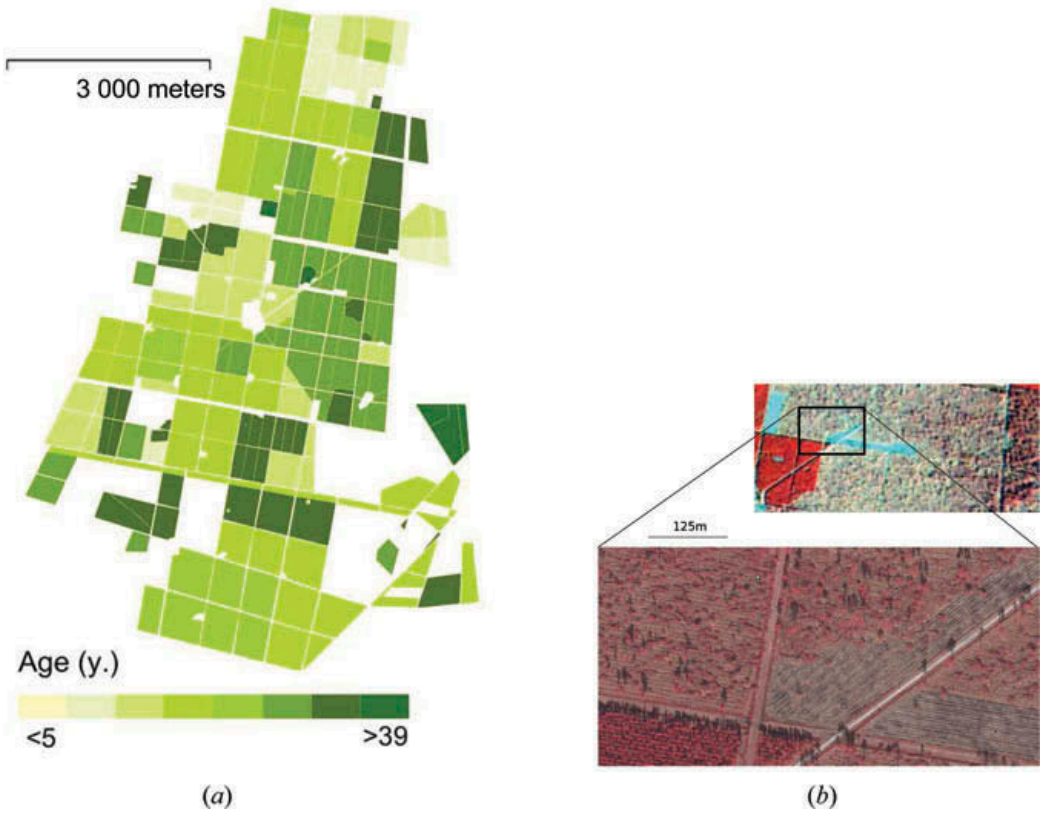


Figure 2. Ancillary data: (*a*) geodatabase of stand ages and (*b*) 15 cm-resolution colour infrared aerial photograph acquired after the storm (26 February 2009).

Downloaded by [Inra] at 00:00 10 July 2014

have 8 m spatial resolution and four spectral bands (blue (B), green (G), red (R), and near-infrared (NIR)). Image radiance was converted into TOA (top-of-atmosphere) reflectance and rescaled between 0 and 255. Both images were orthorectified and georeferenced. Ancillary data include a GIS layer of tree stand delimitations and ages provided by the INRA[1] geodatabase for the Nezer site (Figure 2(*a*)).

In addition, 100 reference areas (damaged vs. non-damaged) were identified on 15 cm-resolution colour infrared photographs acquired after the storm (Figure 2(*b*)). The visual interpretation of these images was complex due to the variety of damage types (bent, fallen, uprooted, and broken-topped trees). Consequently, to minimize damage identification errors, reference areas were identified and delineated as either non-damaged or completely damaged. The size of the reference areas varied between 198 m^2 (around 3 pixels) and 23,128 m^2 (around 260 pixels), with an average of 4806 m^2 (around 75 pixels). All age classes were sampled in both categories (damaged and non-damaged) except for stands younger than five years, in which the damage was limited to bent trees and was difficult to distinguish visually.

The youngest damaged samples (5–13 years) were essentially bent trees. The 14–19-year-old damaged samples were represented by bent, fallen, or uprooted trees. For trees older than 20 years, all damage types were observed (i.e. bent, fallen, uprooted, and broken-topped trees). The oldest damaged samples (>25 years) were primarily groups of uprooted trees. Intact reference areas were also delineated for all age classes. However, intact areas were in the minority for the oldest classes, which were severely damaged by the storm.

3. Methodology

The proposed method is an object-based, bitemporal classification of storm damage. In the following section, selected image features are presented. MS segmentation, which is at the core of our algorithm, will be detailed. Then, the automatic feature selection process will be explained and finally the automatic object-based bitemporal classification method will be presented.

3.1. Proposed input features

Input features can be grouped into three main categories: spectral, textural, and temporal (Table 1). For spectral features, the Formosat-2 spectral bands (B, G, R, NIR) were used as well as two vegetation indices, the normalized difference and soil adjusted vegetation indices (NDVI and SAVI, respectively). First-order statistics, such as the mean and variance of the reflectance, were used as textural features. Among the more commonly used texture features are the Haralick features (derived from a grey level co-occurrence matrix) (Haralick, Shanmugam, and Dinstein 1973), and some of these are considered particularly relevant for forest applications of image analysis (Ruiz, Fdez-sarria, and Recio 2004; Kayitakire, Hamel, and Defourny 2006; St-Louis et al. 2006; Trias-Sanz, Stamon, and Louchet 2008; Tuominen and Pekkarinen 2005). The neighbourhood radius and displacement vector length were both experimentally set to 1 pixel, which corresponds to a maximum crown diameter of 8 m. The displacement vector orientation was horizontal. Finally, three common temporal features were considered: mean correlation, difference, and ratio between the pair of images. Temporal and textural features were processed for each spectral band. A total of 84 features were used: 6 spectral features in addition to 10 textural and 3 temporal, all of the latter computed separately for each of the 6 spectral features.

3.2. MS segmentation

The MS algorithm is a non-parametric, feature-space analysis technique that has shown excellent results in colour image clustering and object delineation (Fukunaga and Hostetler 1975; Comaniciu and Meer 2002). It is based on a density mode searching and clustering technique. The feature space is considered the empirical probability density function (PDF) of the input features. The algorithm proposes a filtering step that associates each pixel in the image with the closest local mode in the density distribution of the feature space. The MS procedure locates these modes without estimating the global density. Then, the fusion of the regions associated with nearby modes leads to image segmentation. The implementation of the method proposed by Comaniciu and Meer (2002) searches for local modes in the joint feature and spatial domain of $n + 2$ dimensions, where n is the number of features added to

Table 1. Spectral, textural, and temporal input features.

Spectral	Textural	Temporal
Blue	Mean	Difference
Green	Variance	Ratio
Red	8 Haralick features	Mean correlation
NIR	Mean, contrast, entropy	
NDVI	Angular second Moment	
SAVI	Inverse difference moment, sum average, sum entropy, sum variance	

Downloaded by [Inra] at 00:00 10 July 2014

the two dimensions of the spatial domain. An iterative procedure of mode seeking consists of shifting the $n + 2$ dimensional window to a local mode.

The search window involves two user-defined inputs that can be deduced from the desired object sizes or physical properties. A radiometric range (h_r) corresponds to the unique spectral radius in the n-dimensional search window and a spatial bandwidth (h_s) that corresponds to the spatial radius of the window. To extract the objects of interest (i.e. tree stands), h_r should be (1) higher than the maximum radiometric difference between intra-region pixel pairs and (2) lower than the radiometric difference between region pixels and surrounding pixels outside the region. In practice, h_r is optimized by the automatic feature selection process (see Section 3.3). All features were individually rescaled between 0 and 255, and radiometric range h_r varied from 2 to 60. The value of h_s should be close to the size of the objects of interest. In fact, the $n + 2$-dimensional moving window for mode searching should remain spatially within the object of interest such that all of its pixels are associated with a local mode within the window. Figure 3 illustrates the impact of both parameters on the segmentation results.

The spatial radius *hs* of the MS segmentation was set using prior thematic knowledge of the desired object size. Using $h_s = 3$, the spatial window can reasonably be considered a regular forest pattern (Orny et al. 2010). Finally, in the proposed method, the MS procedure is used in both segmentation and classification steps using the joint spatio-spectral domain and solely the spectral domain, respectively (see Section 3.4).

3.3. *Automatic feature selection for segmentation*

In the literature, segmentation is usually processed on all available before- and after-image bands (B, G, R, NIR) (Mäkelä and Pekkarinen 2001; Mallinis et al. 2008) or on derived

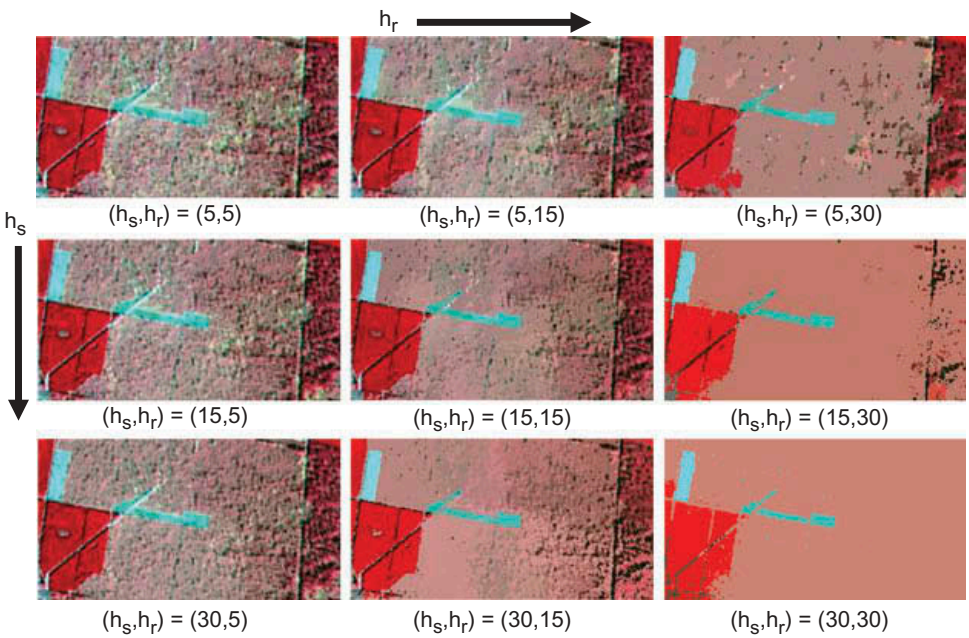


Figure 3. Mean shift segmentations of a four-channel image (B, G, R, NIR) using different parameters (h_s,h_r). When h_r increases, only highly contrasted and homogeneous regions remain (e.g. the intact young stand on the left, the blue farming area, and the medium-aged stand on the right). If h_s increases, only larger regions remain: shadow areas and standing tree groups disappear within the central, strongly damaged area.

Downloaded by [Inra] at 00:00 10 July 2014

bands such as NDVI, SAVI, or transformed bands (PCA) with no feature selection (Wang and Xu 2010). Supervised feature selection for classification can be based on the modelling results of physical variables. In King et al. (2005), the image texture features best correlated to the percentage of crown loss were used for damage classification. In Trias-Sanz, Stamon, and Louchet (2008), a quantitative measure between segmented region boundaries and a reference GIS layer enabled the selection of the best features from among radiometric and texture features.

3.3.1. *Feature selection using test frames*

In this study, to determine the most relevant features for segmentation, input feature selection was carried out through an original generic calibration-like procedure using a test frame. Generally, a test frame is an ensemble of reference images that are used to calibrate cameras. In this study, this process was used to select the best features for segmentation and also aimed at automatically optimizing the segmentation parameters.

Classes describing different forest development stages and structures were defined. The number of classes was denoted as N_c. For the Nezer site, the age GIS reference layer and photointerpretation were used. The test frames were then constructed with N_c small non-overlapping image samples corresponding to N_c classes, yielding as many test frames as input features. The test frames were then segmented by the MS algorithm using one input feature (single test frame) and multiple rescaled features (multiband test frame) while simultaneously testing various segmentation parameterizations. All features were individually rescaled between 0 and 255. The feature segmentation performance (SP) is defined as

$$SP = \frac{1}{nc}\sum_i \frac{\max_j\left[A(R_i)\cap A\left(R_{s_j}\right)\right]}{A(R_i)}, R_i\cap R_{s_j} \neq 0, \tag{1}$$

where A is the area, N_c the number of test frame classes, R_i is a test frame region, and R_{s_j} a segmented region that intersects R_i. i and j are the test frame and segmented region indices, respectively. SP depends on the regions' overlap percentage. Figure 4 illustrates an example of a test frame for after-storm segmentation with four samples and the corresponding SP computation. The highest SP value indicates the best feature or set of

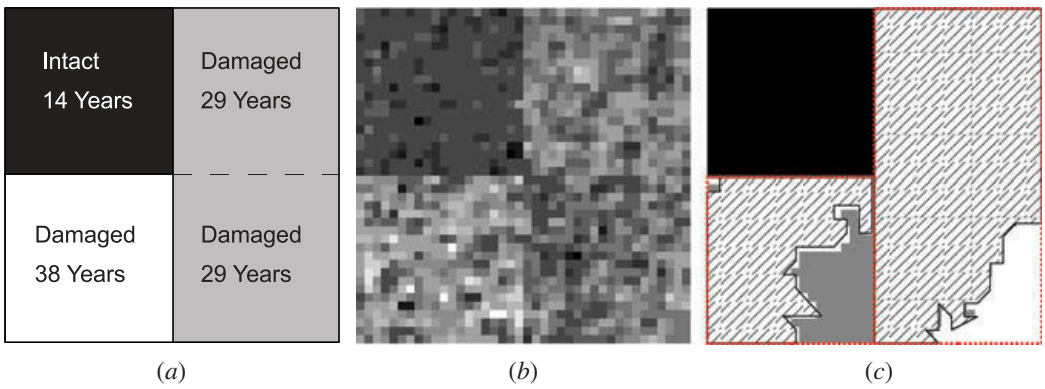


Figure 4. (*a*) Actual regions, (*b*) NIR band test frame, and (*c*) segmented regions using $(h_s,h_r) = (5,10)$. Maximum overlaps per region are hatched. Segmentation performance (SP) = 86%.

Downloaded by [Inra] at 00:00 10 July 2014

features. In this study, test frames were constructed for each feature with four reference regions of 20 × 20 pixels (Figure 6).

Automatic feature selection is applied in three steps: (1) segmentation into structurally homogeneous before-storm regions, (2) segmentation of the after-storm image into temporally and structurally homogeneous regions, and (3) binary change classification. As shown in Figure 6, reference test frames were adapted to each step using (1) actual age classes, (2) classes combining actual age and change type, and (3) actual change classes, respectively. For step (1), temporal features were not used, and for binary classification (step 3), only temporal features averaged over-segmented after-storm regions were used.

SP and the optimal h_r were then processed automatically based on the reference test frames to select the best features for each step.

3.4. Unsupervised object-based bitemporal classification

The global unsupervised object-based bitemporal change detection scheme is depicted in Figure 5. We assume that the degree of change depends not only on the type and intensity of the damage but also on the initial stand structure. Consequently, the before- and after-storm images are segmented independently using the MS algorithm. The before-storm segments correspond to homogeneous structural regions (i.e. those belonging to the same age class). The after-storm segments reflect the degree of change. The selected features are processed as explained in Section 3.3. The joint classification is based on the MS algorithm and is detailed in the current section.

3.4.1. MS spectral classification

The after-storm segmented regions are characterized by object mean temporal descriptors (i.e. by temporal features that are averaged per region). The automatic feature selection process (see Section 3.3) provides the input features that optimize the MS spectral classifier. They are then clustered automatically into change classes using this optimized MS spectral classifier. Unlike MS segmentation (Comaniciu and Meer 2002), this modified version is independent of pixel positions and involves only the spectral domain, which allows the clustering of similarly damaged regions that are spatially distant into the same change class. The MS classifier has a single parameter, h_r, which is the radiometric range of the MS mode-seeking window. The lower the value of h_r, the lower a cluster's intra-variance.

MS spectral classification leads automatically to many change clusters. We did not seek to interpret the significance of these change classes, which would have required collecting specific reference data just after the storm. Our objective in the present study was to produce a damage map with a minimum of field data and an unsupervised classification process. Consequently, in this study, our work was limited to the production of a binary change map even when the MS spectral classification provided multiple change classes. A minimal radiometric range of $h_r = 2$ was used to detect subtle changes (Figure 6). In this context of fine change detection, the number of change clusters can be high, and the spectral distinction between intact and lightly damaged regions may be complex. Thus, the automatic grouping of change clusters into damaged vs. non-damaged classes is challenging. To address this issue, an innovative spatio-temporal feature, the fragmentation rate, is proposed.

Downloaded by [Inra] at 00:00 10 July 2014

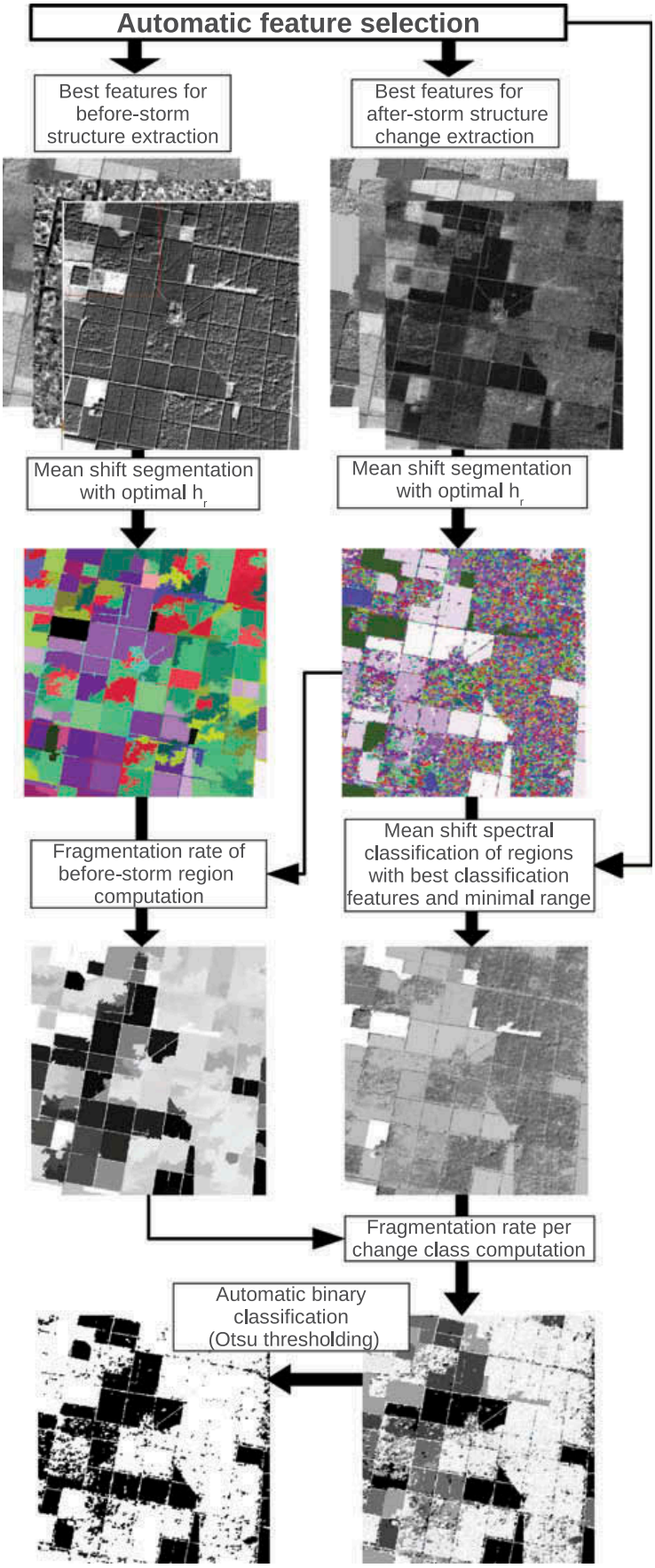


Figure 5. Bitemporal object-based change detection scheme illustrated over a 3 × 3, 3 km^2 area.

Downloaded by [Inra] at 00:00 10 July 2014

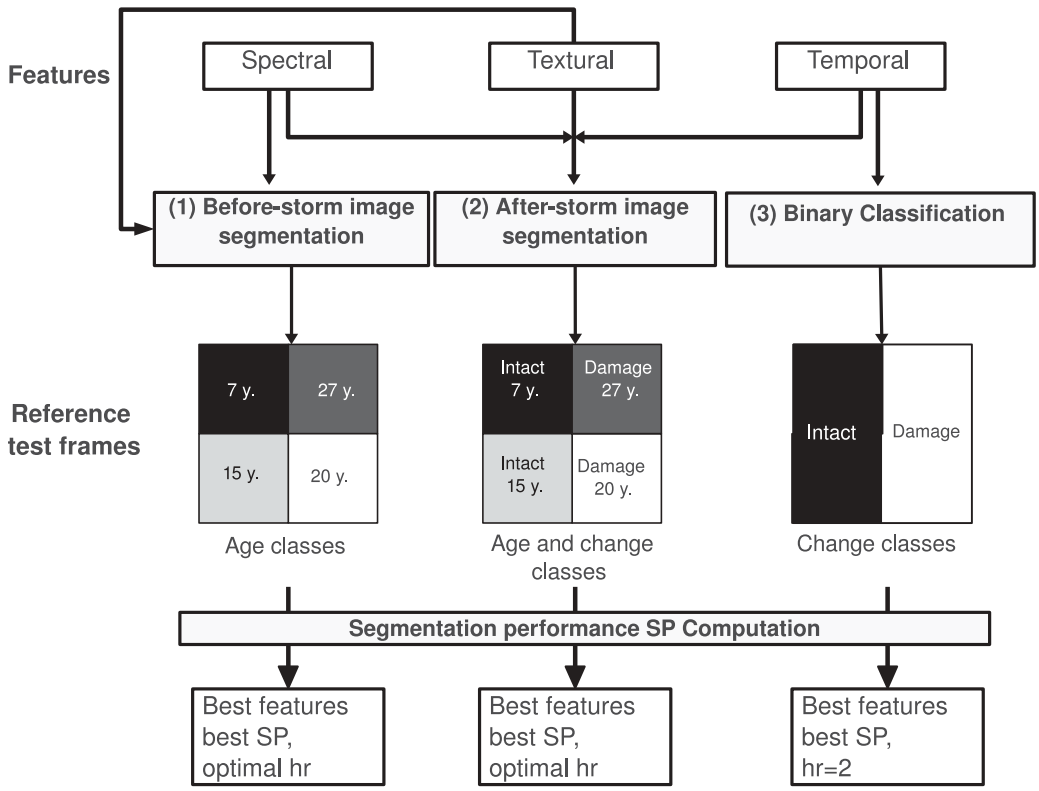


Figure 6. Automatic feature selection process.

Downloaded by [Inra] at 00:00 10 July 2014

3.4.2. *A new spatio-temporal descriptor*

To automatically group the change clusters into damaged vs. non-damaged classes in a robust way, the clusters were characterized by a new spatio-temporal feature: the fragmentation rate (FR). This idea is based on the fact that, generally, for a change detection application, landscape complexity increases with time. The diversification and fragmentation of a landscape lead to more heterogeneous regions in the image. The presence of heterogeneity (more or less important) in an initially homogeneous area can be interpreted as a change. Fragmentation would be more important in the case of disastrous changes such as storms, floods, or earthquakes than for progressive changes such as urbanization. In this study, damaged areas are heterogeneous and therefore appear over-segmented in the after-storm image. Conversely, intact areas correspond to larger regions that have similar delimitations in both images. FR characterizes the before-storm regions and reflects their over-segmentation in the after-storm image. It is computed as a comparison between the before- and corresponding after-storm regions, as follows:

$$FR(R_t) = 1 - \frac{\max_j \left[A(R_t) \cap A\left(R^j_{t+dt}\right) \right]}{A(R_t)}, R_t \cap R^j_{t+dt} \neq 0, \tag{2}$$

where A is the region of interest, R_t is the before-storm region, and R^j_{t+dt} the after-storm regions that are included (partly or entirely) in the before-storm one where j is the after-storm segmented regions index. The more the region R is fragmented, the nearer FR is to 1. Conversely, intact regions should have a low FR close to 0.

The average FR is then computed for the after-storm change clusters. The change cluster FR is then defined as

$$FR(\mathrm{CC}) = \frac{1}{N} \sum_{p \in CC} |[FR(R_t(p))], \tag{3}$$

where CC is the after-storm change cluster and N the number of corresponding pixels (p represents a pixel). The FR per change cluster is the averaged FR of corresponding before-storm regions. The change clusters are finally divided into damaged vs. non-damaged classes, based on their FR and using the unsupervised Otsu threshold (Otsu 1979) which minimizes intra-class variance.

4. Results and discussion

4.1. Feature selection

The segmentation and classification steps for the selected features using mono- and multiband test frames are presented in this section.

4.1.1. Feature selection for segmentation

Table 2 presents the segmentation performance SP (Equation (1)) and MS parametrization for segmentation steps (1) and (2) presented in Section 3.3.1.

One may observe that the best features for segmenting before- and after-storm images are related to the red bands. In fact, as shown by Guyon, Courrier, and Berbigier (2001) for the same Nezer forest site, the pine crown cover fraction, related to tree density and crown size, explains a large part of the variability of red reflectance whereas NIR reflectance depends more strongly on the green biomass of pine trees and undergrowth vegetation. These properties of the red band are important in describing the forest structure before the storm (i.e. the forest variables determining the cover fraction of standing trees) and in detecting their changes after the storm (i.e. the broken, uprooted, or fallen trees that reduce the crown cover fraction without an immediate change in green biomass).

The best feature for the after-storm segmentation is temporal. The information on the initial stand structure (which depends on tree age) from before the storm helps to determine the degree of change after the storm.

Table 2. Optimal features and radiometric ranges, h_r, for after- and before-storm segmentations using a four-region test frame. SP is segmentation performance.

After-storm segmentation			Before-storm segmentation		
Feature	h_r	SP (%)	Feature	h_r	SP (%)
Red ratio	17	87.2	Red	2	80.7
Red	**3**	78.1	NDVI	3	77.3
Green ratio	16	75.5	Green	5	68.7

Downloaded by [Inra] at 00:00 10 July 2014

From Table 2, one may also observe that the maximal SP value is higher when segmenting the after-storm image (87.2%) than the before-storm (80.7%). In fact, the after-storm regions are better discriminated because they combine age class and type of change information, whereas before-storm regions differ only by age classes, which are not easily discriminated.

The optimal features for both segmentations are essentially spectral or temporal but not textural. This is most likely due to, on one hand, the spatial resolution of Formosat-2 (8 m), which provides insufficient textural information in a forest context, and on the other hand, the Haralick parameters not being optimized.

One may observe that the best SP values were obtained using only one feature. Multi-feature segmentation appears to be less relevant in this study, which can be explained by two factors. First, the chosen implementation of MS (Comaniciu and Meer 2002) involves one unique radiometric range, h_r for multiple features. An adaptive MS radiometric range for each feature would enhance the results, as that approach would be more appropriate for remote-sensing images where the reflectance distribution varies for each spectral band. Second, forest canopies are very complex, and changes in their reflectance between the before- and after-storm images are correlated to other changes beyond structural changes following storm damage. For instance, some changes may be detected due to forest phenology or differences in shadows. Using only one feature may limit these detections, considered as misclassifications for the damage class.

Downloaded by [Inra] at 00:00 10 July 2014

4.1.2. Feature selection for classification

The feature selection for the binary classification into damaged vs. non-damaged classes is shown in Table 3. Mono- and multiband two-region test frames were used (see Section 3.3.1). Only temporal features were tested (see Section 3.1). The best-performing feature was the green band difference, with a high SP value of over 92.6%. This was therefore used as an input for the MS classifier to establish the change classes (see Section 4.4). As with feature selection for segmentation, multiple features did not provide better results.

4.2. MS segmentation

Figure 7 depicts the segmentation results before and after the storm using the corresponding best features. One can visually distinguish the intact and damaged areas. Intact areas are larger and have similar delimitations in both segmentations, whereas the damaged areas are more heterogeneous, leading to over-segmented regions.

Table 3. Optimal features for binary classification using a two-region test frame with $h_r = 2$.

Feature	SP (%)
Green difference	92.6
NIR difference	70.5
Red difference	64.5

(*a*) (*b*)

Figure 7. Segmentation of before- and after-storm images using the best features (i.e. red band and red band ratio, respectively).

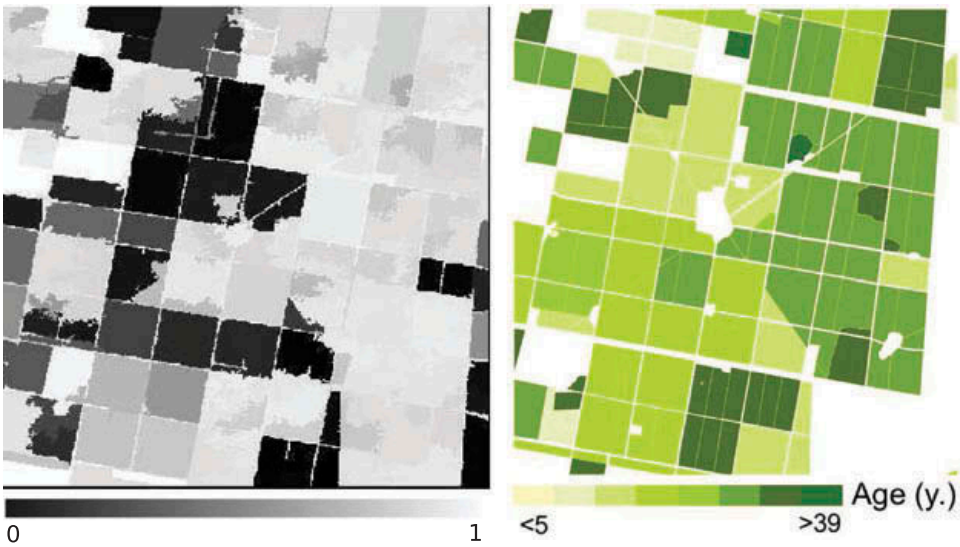
Downloaded by [Inra] at 00:00 10 July 2014

4.3. FR

Figure 8 illustrates the FR of before-storm segments in grey levels. The lighter the regions, the more damaged they are. This result visually matches the tree stand age map, where older stands appear to have more extensive damage than younger stands. In fact, among numerous factors, tree height most strongly influences the sensitivity of a tree to the wind (Cucchi et al. 2005). The young stands are dense with small trees, which makes them more robust to the wind. In contrast, older stands are less dense and more heterogeneous due to silvicultural practices and are composed of taller trees that are more vulnerable and likely to be damaged by storms.

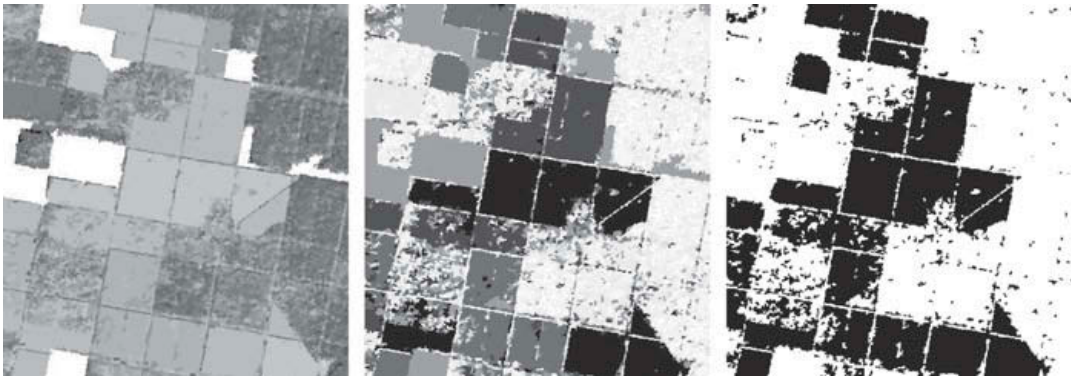
4.4. MS binary classification

The MS classifier is based exclusively on the spectral domain. After-storm segmented regions are classified into change clusters. Figure 9(*a*) shows approximately 30 change clusters



(*a*) (*b*)

Figure 8. Comparison between fragmentation rate and tree stand ages.



(a) (b) (c)

Figure 9. MS spectral classification and fragmentation rate for automatic binary mapping of forest storm damage.

created using the MS spectral classifier. Given the histogram complexity of this image, it is hard to automatically separate the obtained clusters into damaged vs. non-damaged classes. By contrast, one may observe (Figure 9(*b*)) that after characterization of these regions by their average FR (3)), the damaged areas are better discriminated. Figure 9(*c*) shows the result of an automatic binary classification achieved using Otsu thresholding.

4.5. Map validation

4.5.1. Comparison to a pixel-based approach

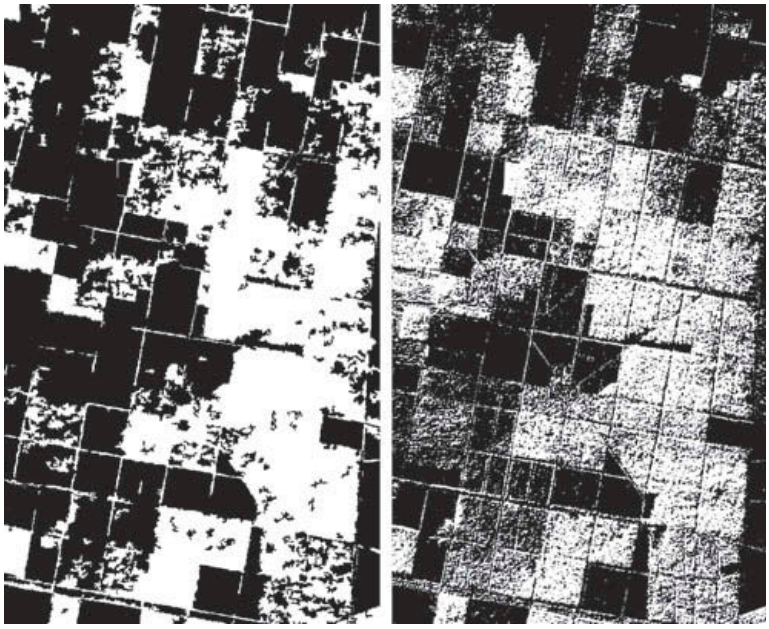
Our classification method was compared to an unsupervised pixel-based approach based on the well-known *k*-means clustering algorithm (Duda, Hart, and Stork 2001) and using all available image bands before and after the storm (eight bands) (i.e. without any feature selection process). The temporal feature used was the band difference, as our feature selection process indicated its efficiency for binary classification (see Section 4.1.2).

Table 4 shows the global confusion matrices obtained by comparing the pixel values from the classification results to reference data for both methods. The object-based

Table 4. Confusion matrices for binary classification (damage vs. non-damage) using our object-based approach with feature selection and a K-means pixel-based approach without feature selection.

	Ground truth			
	Undamaged	Damaged	Total	Commission error (%)
Object-based approach				
Undamaged	1436	459	1895	24.22
Damaged	81	2452	2533	3.2
Total	1517	2911	4428	
Omission error (%)	5.34	15.77		
Overall accuracy (%)	87.80			
Pixel-based approach				
Undamaged	1361	788	2149	36.67
Damaged	156	2123	2279	6.85
Total	1517	2911	4428	
Omission error (%)	10.28	27.07		
Overall accuracy (%)	78.68			

Downloaded by [Inra] at 00:00 10 July 2014



(a) (b)

Figure 10. Change detection maps. Damaged vs. non-damaged areas are depicted in white and black, respectively.

Downloaded by [Inra] at 00:00 10 July 2014

approach outperformed the pixel-based approach with an overall classification accuracy of 87.8% vs. 78.7%, respectively. This result confirms the relevance of automatic feature selection and the new fragmentation descriptor to the proposed change detection scheme.

Additionally, one may observe that omission and commission errors were twofold less important for our method than for the pixel-based method. Few omission errors were obtained in non-damaged regions (5.3%), and few commission errors were found in damaged regions (3.2%). The final binary maps exhibiting damaged vs. non-damaged areas for both methods are depicted on Figure 10. As expected, due to both the optimized segmentation step and the spatio-temporal descriptor introduced, the object-based approach leads to a smoother change detection map.

4.6. Influence of feature selection on classification accuracy

To assess the relevance of our feature selection method, the segmentation of the bi-temporal image was processed using the eight initial bands (i.e. before- and after-storm images) with no feature selection. The overall classification accuracy decreased from 87.8 to 83.3% when using feature selection.

For the K-means pixel-based approach, the best spectral feature led to an overall classification accuracy of 83.2%, which is better than that using the eight initial bands (78.7%). This confirms, in accordance with our object-based approach (see Section 4.1.1), that the use of all bands is not necessarily the best choice for bitemporal segmentation and classification and that a feature selection process improves the classification results.

4.6.1. Comparison to age class map

The INRA geodatabase, which collects the ages of all pine stands, was compared to the obtained classification. The accuracy assessment is detailed according to age class in Table 5.

Table 5. Confusion matrix according to age class.

Age class	Number of pixels	Omission error (%)		Commission error (%)		Overall accuracy per age class (%)
		Non-damage	Damage	Non-damage	Damage	
5 − 9	115	0	100	40.0	100	60.0
10 − 13	510	0	99.0	59.0	100	41.4
14 − 20	1406	0.76	12.0	11.9	0.76	93.3
21 − 25	424	6.25	0	0	10.1	96.0
26 − 30	694	9.30	0	0	0.61	99.4
31 − 39	1086	26.3	0	0	5.53	95.2
>39	193	11.1	30.3	76.8	1.61	71.5

The percentage of pixels classified as damaged (damage rate) was computed within each age class and the results are shown in Figure 11, which also depicts the behaviour of the overall classification accuracy with respect to each age class. This figure shows that classification accuracy increases significantly with stand age. In fact, tree sensitivity to wind increases with tree height, which depends on tree age (Cucchi et al. 2005). This effect is confirmed by the Formosat-2-derived damage rate, which was higher for older stands and reached 70% for stands over 25 years old. This is a good indicator of the quality of the results obtained.

Table 5 shows that the classification accuracy of stands aged from 14 to 39 years (four intermediate classes) is high and ranges between 93.3 and 99.4%. However, the two youngest classes and the oldest stand class have lower detection rates (60, 41.4, and 71.5%, respectively) resulting from high omission and commission errors (>30%). Confusion occurred between damaged areas and non-damaged older areas, which were originally heterogeneous and sparse. In fact, both areas have similar spectral properties and textural structure. In addition, subtle change areas with minor tree damage (bent trees) were hard to detect in dense young stands. In addition, very small intact or damaged areas cannot be detected due to the limited spatial resolution of Formosat-2 images. Finally, in our method, shadows are not explicitly taken into account in segmentation or in

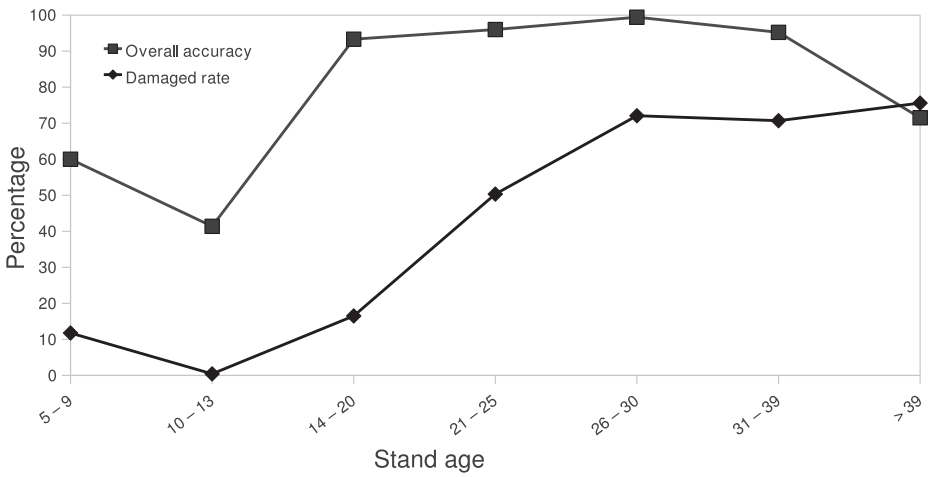


Figure 11. Overall classification accuracy and damage rate as a function of age class.

Downloaded by [Inra] at 00:00 10 July 2014

classification, potentially leading to misclassifications. With only one image dataset, we were unable to estimate the influence of shadows or viewing angle on the outcome of the change detection. These issues should be further developed in future work.

5. Conclusion

We have presented in this paper an object-based bitemporal change detection method that is well suited to emergency mapping. It has the appealing property of requiring only a few samples to construct the test frames, leading to a low requirement for supervision compared with more traditional supervised methods. Moreover, our method involves only two user-defined parameters and does not need to meet any statistical assumptions, thanks to the powerful MS clustering algorithm at the core of our change detection scheme.

Our contribution comprises two main novelties with respect to similar works in forestry. First, an automatic feature selection process, applied to both the segmentation and classification steps, is introduced. Its originality lies in the use of test frames (single- or multibands) of adequate forest samples. This innovative feature selection process, inspired by camera calibration procedures, allows a rapid evaluation of hundreds of features and combined features. It is applicable to the optimization of any other segmentation algorithm.

The second original feature of our approach is a relevant spatio-temporal descriptor: the FR, which enables a simple automatic binary threshold (damage vs. non-damage) for the various change clusters obtained by the MS classifier. The FR can be applied to other change detection applications assuming that natural processes, either progressive (urbanization) or disastrous (floods, earthquakes, etc.), tend to fragment the landscape.

Our change detection scheme was applied in forestry, in the context of storm damage mapping. Severe damage was accurately detected, and our method yields a good overall accuracy of 87.8

In future work, the radiometric range, h_r should be adapted to each feature in the MS algorithm to cover the spectral distribution variability and therefore take advantage of the combination of multiple features. In addition, our MS-based change detection scheme allows the detection of gradual changes. Appropriate reference data should be collected to validate such applications. In addition, our method will be applied to very high-resolution satellite images, such as those of QuickBird, IKONOS, or Pleiade, to evaluate its performance for subtle change detection (e.g. bent trees) and other applications.

Acknowledgements

The Orfeo programme facilitated the future use of the VHR Pleiades images. Access to the images was granted by the CNES Kalideos Programme (http://Kalideos.cnes.fr). The images were analysed using the OTB toolbox (www.orfeo-toolbox.org). The authors also acknowledge Olivier Hagolle (CNES/Cesbio, Toulouse), and Christophe Moisy (INRA, EPHYSE, Bordeaux) for image preprocessing.

Funding

The present study was carried out within the Orfeo programme funded by CNES (Centre National d'Etudes Spatiales). The authors also acknowledge INRA for the funding of the orthophotographs within *Crédit tempête 2009.*

Downloaded by [Inra] at 00:00 10 July 2014

Note

1. Institut National de Recherches Agronomiques (French National Institute of Agricultural research).

References

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. 2004. "Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for Gis-Ready Information." *ISPRS Journal of Photogrammetry and Remote Sensing* 58 (3–4): 239–258. doi:10.1016/j.isprsjprs.2003.10.002.

Blaschke, T. 2010. "Object Based Image Analysis for Remote Sensing." *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1): 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Chehata, N., C. Orny, S. Boukir, and D. Guyon. 2011. "Object-Based Forest Change Detection Using High Resolution Satellite Images." In *PIA11, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII*, edited by U. Stilla, F. Rottensteiner, H. Mayer, B. Jutzi, and M. Butenuth, Vol. 3/W22, Munich, October 5–7, 49–55.

Clandillon, S., H. Yesou, and C. Meyer. 2003. "Benefits of SPOT 5 HR and VHR Data for Forest Management and Windfall Damage Mapping." In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS '03*, Vol. 5, Toulouse, 3079–3081. IEEE. doi:10.1109/IGARSS.2003.1294689.

Comaniciu, D., and P. Meer. 2002. "Mean Shift: Arobust Approach toward Feature Space Analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (5): 603–619. doi:10.1109/34.1000236.

Conchedda, G., L. Durieux, and P. Mayaux 2008. "An Object-Based Method for Mapping and Change Analysis in Mangrove Ecosystems." *ISPRS Journal of Photogrammetry and Remote Sensing* 63 (5): 578–589. Remote Sensing of the Coastal Ecosystems Special Issue. doi:10.1016/j.isprsjprs.2008.04.002.

Cucchi, V., C. Meredieu, A. Stokes, F. Coligny, J. Suarez, and B. A. Gardiner. 2005. "Modelling the Windthrow Risk for Simulated Forest Stands of Maritime Pine (Pinus Pinaster Ait.)." *Forest Ecology and Management* 213: 184–196. doi:10.1016/j.foreco.2005.03.019.

Desclée, B., P. Bogaert, and P. Defourny. 2006. "Forest Change Detection by Statistical Object-Based Method." *Remote Sensing of Environment* 102 (1–2): 1–11. doi:10.1016/j.rse.2006.01.013.

Duda, R. O., P. E. Hart, and D. G. Stork. 2001. *Pattern Classification*. 2nd ed. New York: Wiley-Interscience.

Franklin, S. E., R. H. Waring, R. W. McCreight, W. B. Cohen, and M. Fiorella. 1995. "Aerial and Satellite Sensor Detection and Classification of Western Spruce Budworm Defoliation in a Subalpine Forest." *Canadian Journal of Remote Sensing* 21: 299–308.

Fraser, R. H., A. Abuelgasim, and R. Latifovic. 2005. "A Method for Detecting Large-Scale Forest Cover Change Using Coarse Spatial Resolution Imagery." *Remote Sensing of Environment* 95 (4): 414–427. doi:10.1016/j.rse.2004.12.014.

Fukunaga, K., and L. Hostetler. 1975. "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition." *IEEE Transactions on Information Theory* 21: 32–40. doi:10.1109/TIT.1975.1055330.

Garestier, F., P. Dubois-Fernandez, D. Guyon, and T. Le Toan. 2009. "Forest Biophysical Parameter Estimation Using L- and P-Band Polarimetric SAR Data." *IEEE Transactions on Geoscience and Remote Sensing* 47 (10): 3379–3388. doi:10.1109/TGRS.2009.2022947.

Guyon, D., I. Champion, C. Meredieu, A. Kruszewski, and T. Labbe. 2005. *Final Report – Work Package 1-1* "*Ground Data Base* ", April 2005. Technical Report, INRA/Egyptsar04/IC01 and ESTEC-European Space Agency, Contract 18507/04, 18.

Guyon, D., G. Courrier, and P. Berbigier. 2001. "Sensitivity Analysis of Coniferous Forest Reflectance with Canopy Structure and Undergrowth Characteristics from Satellite Data and Modelling (Case Study: Landes Maritime Pine Forest)." In *8th International Symposium* "*Physical Measurements and Signatures in Remote Sensing*", edited by CNES, Aussois, January, 387–392.

Downloaded by [Inra] at 00:00 10 July 2014

Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man, and Cybernetics* 3: 610–621. doi:10.1109/TSMC.1973.4309314.

Hayes, D. J., and W. B. Cohen. 2007. "Spatial, Spectral and Temporal Patterns of Tropical Forest Cover Change as Observed with Multiple Scales of Optical Satellite Data." *Remote Sensing of Environment* 106 (1): 1–16. doi:10.1016/j.rse.2006.07.002.

Huang, C., K. Song, S. Kim, J. R. G. Townshend, P. Davis, J. G. Masek, and S. N. Goward. 2008. "Use of a Dark Object Concept and Support Vector Machines to Automate Forest Cover Change Analysis." *Remote Sensing of Environment* 112 (3): 970–985. doi:10.1016/j.rse.2007.07.023.

Im, J., and J. R. Jensen. 2005. "A Change Detection Model Based on Neighborhood Correlation Image Analysis and Decision Tree Classification." *Remote Sensing of Environment* 99 (3): 326–340. doi:10.1016/j.rse.2005.09.008.

Johansen, K., L. A. Arroyo, S. Phinn, C. Witte, R. J. Kauth, and G. S. Thomas. 2010. "Comparison of Geo-Object Based and Pixel-Based Change Detection of Riparian Environments Using High Spatial Resolution Multi-Spectral Imagery." *Photogrammetric Engineering and Remote Sensing* 76 (2): 123–136. doi:10.14358/PERS.76.2.123.

Johansen, K., N. C. Coops, S. E. Gergel, and Y. Stange. 2007. "Application of High Spatial Resolution Satellite Imagery for Riparian and Forest Ecosystem Classification." *Remote Sensing of Environment* 110 (1): 29–44. doi:10.1016/j.rse.2007.02.014.

Jonikavičius, D., and G. Mozgeris. 2013. "Rapid Assessment of Wind Storm-Caused Forest Damage Using Satellite Images and Stand-Wise Forest Inventory Data." *iforest* 6: 150–155. doi:10.3832/ifor0715-006.

Kayitakire, F., C. Hamel, and P. Defourny. 2006. "Retrieving Forest Structure Variables Based on Image Texture Analysis and IKONOS-2 Imagery." *Remote Sensing of Environment* 102: 390–401. doi:10.1016/j.rse.2006.02.022.

Kennedy, R. E., P. A. Townsend, J. E. Gross, W. B. Cohen, P. Bolstad, Y. Q. Wang, and P. Adams. 2009. "Remote Sensing Change Detection Tools for Natural Resource Managers: Understanding Concepts and Tradeoffs in the Design of Landscape Monitoring Projects." *Remote Sensing of Environment* 113 (7): 1382–1396. doi:10.1016/j.rse.2008.07.018.

Kim, M., M. Madden, and T. Warner. 2008. "Estimation of Optimal Image Object Size for the Segmentation of Forest Stands with Multispectral IKONOS Imagery." In *Object-Based Image Analysis*, edited by T. Blaschke, S. Lang, and G. J. Hay, 291–307. Lecture Notes in Geoinformation and Cartography. Berlin: Springer.

King, D. J., I. Olthof, P. K. E. Pellikka, E. D. Seed, and C. Butson. 2005. "Modelling and Mapping Damage to Forests from an Ice Storm Using Remote Sensing and Environmental Data." *Natural Hazards* 35: 321–342. doi:10.1007/s11069-004-1795-4.

Lamonaca, A., P. Corona, and A. Barbati. 2008. "Exploring Forest Structural Complexity by Multi-Scale Segmentation of VHR Imagery." *Remote Sensing of Environment* 112 (6): 2839–2849. doi:10.1016/j.rse.2008.01.017.

Liu, D., K. Song, J. R. G. Townshend, and P. Gong. 2008. "Using Local Transition Probability Models in Markov Random Fields for Forest Change Detection." *Remote Sensing of Environment* 112 (5): 2222–2231. Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue. doi:10.1016/j.rse.2007.10.002.

Mäkelä, H., and A. Pekkarinen. 2001. "Estimation of Timber Volume at the Sample Plot Level by Means of Image Segmentation and Landsat TM Imagery." *Remote Sensing of Environment* 77: 66–75. doi:10.1016/S0034-4257(01)00194-8.

Mallinis, G., N. Koutsias, M. Tsakiri-Strati, and M. Karteris. 2008. "Object-Based Classification Using Quickbird Imagery for Delineating Forest Vegetation Polygons in a Mediterranean Test Site." *ISPRS Journal of Photogrammetry and Remote Sensing* 63 (2): 237–250. doi:10.1016/j.isprsjprs.2007.08.007.

McRoberts, R. E., and B. F. Walters. 2012. "Statistical Inference for Remote Sensing-Based Estimates of Net Deforestation." *Remote Sensing of Environment* 124: 394–401. doi:10.1016/j.rse.2012.05.011.

Meyer, C., P. Geldreich, and H. Yesou. 2001. "Apports des Données Simulées SPOT5 Pour l' Évaluation Des Dégâts De Tempête Dans La Forêt. Cas De La Foret De Haguenau (Alsace, France)." In *Bulletin de la Société Française de Photogrammétrie et de Télédétection*, No. 164–165 (2001-4/2002-1): 162–173.

Downloaded by [Inra] at 00:00 10 July 2014

Olthof, I., D. J. King, and R. A. Lautenschlager. 2004. "Mapping Deciduous Forest Ice Storm Damage Using Landsat and Environmental Data." *Remote Sensing of Environment* 89: 484–496. doi:10.1016/j.rse.2003.11.010.

Orny, C., N. Chehata, S. Boukir, and D. Guyon. 2010. *Characterization of Maritime Pine Forest Structure Changes with VHR Satellite Imagery: Application to the 24th January 2009 Windfall Damages Cartography.* Technical Report. Toulouse: CNES.

Otsu, N. 1979. "A Threshold Selection Method from Gray-Level Histograms." *IEEE Transactions on Systems, Man and Cybernetics* 9: 62–66. doi:10.1109/TSMC.1979.4310076.

Ozdarici, O. A., and Z. Akyurek. 2012. "A Segment-Based Approach to Classify Agricultural Lands by Using Multi-Temporal Optical and Microwave Data." *International Journal of Remote Sensing* 33 (22): 7184–7204. doi:10.1080/01431161.2012.700423.

Pasher, J., and D. J. King. 2009. "Mapping Dead Wood Distribution in a Temperate Hardwood Forest Using High Resolution Airborne Imagery." *Forest Ecology and Management* 258: 1536–1548. doi:10.1016/j.foreco.2009.07.009.

Ruiz, L. A., A. Fdez-sarria, and J. A. Recio. 2004. "Texture Feature Extraction for Classification of Remote Sensing Data Using Wavelet Decomposition: A Comparative Study." *International Archives of Photogrammetry and Remote Sensing* XXXV: 1682–1750.

Schwarz, M., C. Steinmeier, and L. Waser. 2001. "Detection of Storm Losses in Alpine Forest Areas by Different Methodic Approaches Using High Resolution Satellite Data." In *Proceedings of the 21st EARSeL Symposium: Observing Our Environment from Space: New Solutions for a New Millenium*, 251–257.

St-Louis, V., A. M. Pidgeon, V. C. Radeloff, T. J. Hawbaker, and M. K. Clayton. 2006. "Highresolution Image Texture as a Predictor of Bird Species Richness." *Remote Sensing of Environment* 105: 299–312. doi:10.1016/j.rse.2006.07.003.

Trias-Sanz, R., G. Stamon, and J. Louchet. 2008. "Using Colour, Texture, and Hierarchial Segmentation for High-Resolution Remote Sensing." *ISPRS Journal of Photogrammetry and Remote Sensing* 63: 156–168. doi:10.1016/j.isprsjprs.2007.08.005.

Tuominen, S., and A. Pekkarinen. 2005. "Performance of Different Spectral and Textural Aerial Photograph Features in Multi-Source Forest Inventory." *Remote Sensing of Environment* 94: 256–268. doi:10.1016/j.rse.2004.10.001.

Vibrans, A. C., R. E. McRoberts, P. Moser, and A. L. Nicoletti. 2013. "Using Satellite Image-Based Maps and Ground Inventory Data to Estimate the Area of the Remaining Atlantic Forest in the Brazilian State of Santa Catarina." *Remote Sensing of Environment* 130: 87–95. doi:10.1016/j.rse.2012.10.023.

Wang, W., J. J. Qu, X. Hao, Y. Liu, and J. A. Stanturf. 2010. "Post-Hurricane Forest Damage Assessment Using Satellite Remote Sensing." *Agricultural and Forest Meteorology* 150: 122–132. doi:10.1016/j.agrformet.2009.09.009.

Wang, Y., Y. S. Soh, and H. Shultz. 2006. "Individual Tree Crown Segmentation in Aerial Forestry Images by Mean Shift Clustering and Graph-Based Cluster Merging." *International Journal of Computer Science and Network Security* 6 (11): 40–45.

Wang, F., and Y. J. Xu. 2010. "Comparison of Remote Sensing Change Detection Techniques for Assessing Hurricane Damage to Forests." *Environmental Monitoring and Assessment* 162: 311–326. doi:10.1007/s10661-009-0798-8.

Wulder, M. A., J. C. White, N. C. Coops, and C. R. Butson. 2008. "Multi-Temporal Analysis of High Spatial Resolution Imagery for Disturbance Monitoring." *Remote Sensing of Environment* 112 (6): 2729–2740. doi:10.1016/j.rse.2008.01.010.

Downloaded by [Inra] at 00:00 10 July 2014

5.4 Axe segmentation

Détection de bâtiments à partir d'une image satellitaire par combinaison d'approches ascendante et descendante

Mohamed Mahmoud Sidi Youssef[1] Clément Mallet[2]
Nesrine Chehata[3] Arnaud Le Bris[2] Adrien Gressin[2]

[1]Sup'Com Ecole Supérieure des Communications de Tunis, Tunisie
[2]IGN/Laboratoire MATIS, Université Paris Est, France
[3]Laboratoire G&E (EA 4592), IPB / Université de Bordeaux, France
prenom.nom@ign.fr

Résumé

La détection de bâtiments à partir d'imagerie géospatiale a été un sujet de recherche très populaire ces vingt dernières années et en particulier avec l'émergence des satellites à très haute résolution. Les méthodes existantes présentent diverses limitations les empêchant d'être efficaces à grandes échelles spatiales et temporelles : elles dépendent du contexte, d'un paramétrage fastidieux ou nécessitent plusieurs sources de données. Dans ce travail, nous proposons une méthode automatique qui répond en partie à ces problèmes en combinant les avantages des méthodes descendantes et ascendantes d'extraction de bâtiments, i.e., à la fois des algorithmes de classification et de reconnaissance des formes. Cela permet de détecter correctement les objets grâce à des connaissances géométriques a priori tout en délimitant de façon précise leurs frontières et en préservant leur forme. La méthode est évaluée sur une zone semi-urbaine de plus de 230 bâtiments en utilisant une image simulée multispectrale Pléiades de 0,5 m de résolution.

Mots Clef

Segmentation, classification, processus ponctuels marqués, bâtiments, imagerie géospatiale à très haute résolution.

Abstract

Building detection for geospatial optical images has been a popular topic of research for the last twenty years and in particular with the emergence of very high resolution satellite images. Existing methods exhibit various flaws and prevent them from being efficient at large scales of space and time : they are context-dependent, require a tedious parameter tuning or several data sources. Therefore, we propose a fully automatic method that alleviates these issues by combining the strengths of bottom-up and top-down approaches, i.e., of both classification and pattern recognition algorithms. This allows to correctly detect the objects by geometric prior knowledge while finely delineating their borders and preserving their shape. The method is evaluated over a complex area of more than 230 buildings using a 0.5 m multispectral pansharpened Pleiades image.

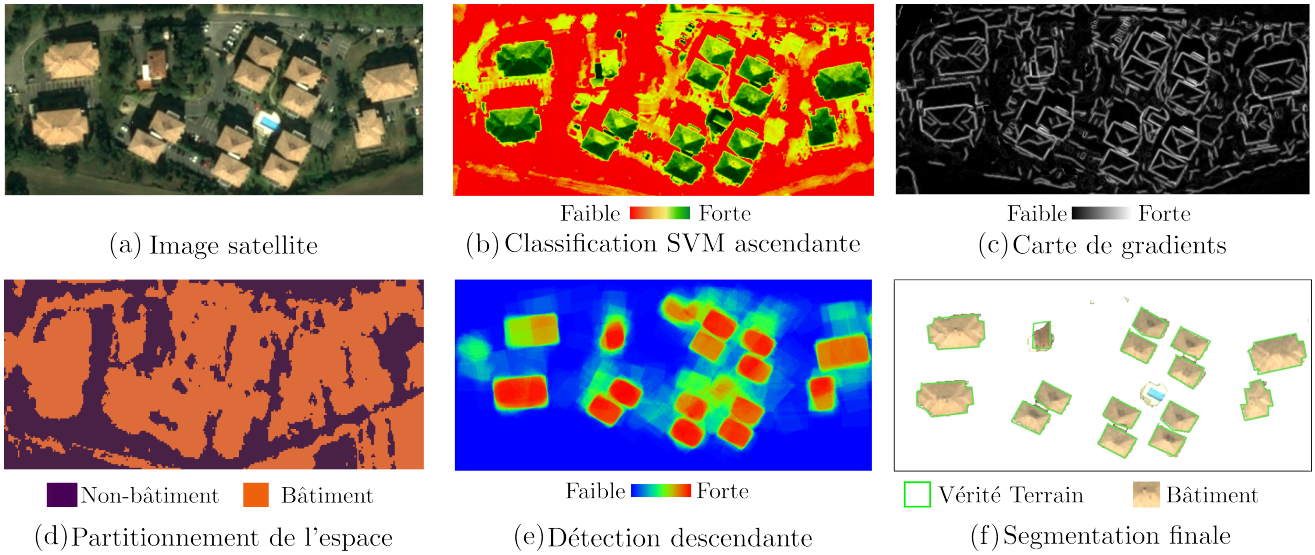
Keywords

Segmentation, classification, Marked point process, building, very high resolution imagery.

1 Introduction

La détection de bâtiments à partir d'images satellitaires à très haute résolution spatiale (THR) est un enjeu important pour de nombreuses applications principalement liées à la mise en place des politiques publiques. Durant les deux dernières décennies, elle a fait l'objet de nombreuses recherches utilisant différentes sources de données telles que les images optiques multispectrales ou hyperspectrales, les données SAR ou encore LIDAR, ainsi que des connaissances externes : bases de données topographiques et cadastre. La source de données la plus simple et la plus utilisée reste les images optiques monoculaires [1] même si de nombreux travaux introduisent l'information d'altitude [2, 3] ou radar [4]. Cependant, de telles données ne sont pas toujours disponibles et le fait d'utiliser différentes sources de données rend ces méthodes moins facilement applicables à grande échelle. Par conséquent, notre étude se focalise sur le traitement d'une unique image satellitaire optique THR monoculaire. Cela offre un bon compromis entre résolutions spatiale (forte), spectrale (faible) et temporelle (moyenne).

Les travaux antérieurs se regroupent en trois catégories. (1) Les méthodes ascendantes (*Bottom-Up*) se fondent sur l'extraction d'information bas-niveau. Les images optiques submétriques fournissent une description détaillée des scènes urbaines d'un point de vue spatial et spectral, permettant l'utilisation simple d'algorithmes de classification supervisée [5, 6]. Cependant, sans modélisation sous-jacente des bâtiments, de telles classifications peuvent êtres très bruitées en raison de la forte hétérogénéité et complexité des scènes urbaines. (2) Les méthodes descendantes (*Top-Down*) se focalisent directement sur la détection d'objets, par reconnaissance de formes et en utilisant des connaissances a priori sur les arrangements spatiaux et les in-



(a) Image satellite (b) Classification SVM ascendante (c) Carte de gradients

(d) Partitionnement de l'espace (e) Détection descendante (f) Segmentation finale

FIGURE 1 – Aperçu général de l'approche (*Zone 1*). Voir le texte pour plus de détails.

teractions entres objets [7, 8]. Cependant, elles sont souvent complexes, hautement paramétrées et avec des temps de calcul rédhibitoires pour des grandes zones. (3) Les méthodes mixtes visent à bénéficier des avantages des deux approches. En général, les approches fondées sur l'optimisation par graphes sont choisies de manière à propager l'information locale obtenue par classification à de grandes zones par un raisonnement spatial structuré [1]. Elles sont efficaces, mais la recherche de la configuration optimale du bâtiment nécessite une optimisation globale difficile à obtenir sur de grandes échelles. Par conséquent, une autre solution est d'incorporer une étape de reconnaissance de formes dans une segmentation d'images [9] ou d'alterner reconnaissance ascendante et détection descendante pour optimiser les deux approches [10, 11, 12]. En télédétection, seuls [13] ont fusionné des méthodes de classification et de reconnaissance de formes pour la détection de bâtiments.
Dans cet article, nous proposons une approche combinant des approches ascendante et descendante pour la détection des bâtiments. Les principales contributions de notre méthode sont les suivantes : une méthode entièrement automatique ne nécessitant qu'une image optique monoculaire, une approche mixte qui combine les avantages des approches ascendante et descendante, et un algorithme indépendant du contexte qui peut être appliqué à de grandes échelles et permet de gérer des scènes avec différentes tailles et configurations de bâtiments.

2 Méthodologie

L'approche proposée est illustrée sur la figure 1, et se décompose en trois étapes principales.

1. Approche ascendante : elle exploite les informations bas-niveau de l'image (section 3). D'une part, une classification supervisée par pixel est effectuée et permet de calculer une carte de probabilité de bâtiments (Figure 1b). D'autre part, des primitives simples (lignes et régions) sont extraites de manière à contraindre, améliorer et accélérer la détection descendante de motifs rectangulaires (Figure 1c).
2. Approche descendante : elle vise également à fournir une carte de probabilité de bâtiments. Nous proposons d'utiliser une modélisation par processus ponctuels marqués (PPM) qui permet de prendre en compte les contraintes géométriques et les interactions spatiales directement au niveau des objets. La carte de probabilité de bâtiments est obtenue par la fusion des extractions de rectangles par accumulation de PPM (section 4) à différents paramétrages (Figure 1e). En effet, puisque les bâtiments peuvent avoir différentes tailles et formes, et sachant que les motifs rectangulaires ne permettent de trouver que l'emplacement grossier des bâtiments, plusieurs hypothèses sont formulées et permettent de mieux délimiter les zones de bâti.
3. Etape de décision : elle fusionne la classification et le partitionnement ascendants avec la détection descendante de motifs rectangulaires. Une segmentation binaire bâtiment Vs. Non-bâtiment est calculée en utilisant une formulation Markovienne, optimisée par une coupure de graphe (Figure 1f).

Notre approche est très similaire au travail décrit dans [13] avec certaines différences : les étapes de classification et d'extraction de rectangles y sont effectuées alternativement jusqu'à convergence de la classification. En pratique, seuls quelques pixels changent de label après quelques itérations. Dans notre article, nous considérons que l'étape ascendante est suffisamment fiable pour effectuer une unique étape d'affinage par approche descendante. En outre, la succession des étapes d'optimisation entre les deux approches est remplacée par l'accumulation de plusieurs cartes de motifs rectangulaires binaires.

3 Approche ascendante

3.1 Classification supervisée

La première étape consiste, pour chaque pixel de l'image, à calculer la probabilité d'appartenance à la classe *bâtiment* (Figure 1b). Aucune analyse orientée-objet n'est réalisée. Le but de cette étape est de simplement mettre en évidence les contours de bâtiments qui ne peuvent pas être détectés de façon dense par l'approche descendante. Aucune régularisation spatiale n'est nécessaire à cette étape, elle sera effectuée lors de l'approche mixte par régularisation Markovienne.

Une approche supervisée est adoptée afin de mieux saisir la variabilité des matériaux de toiture sur l'image, même si une unique classe *bâtiment* est conservée. Une base de données 2D existante de bâtiments est utilisée à cette fin. L'ensemble d'apprentissage est obtenu automatiquement en sélectionnant d'abord au hasard 1000 échantillons de la classe *bâtiment* dans la base de données. Enfin, un classifieur SVM, avec un noyau gaussien est utilisé et permet de générer une carte de probabilité de bâtiments [14]. Les quatre bandes spectrales de l'image satellitaire (*bleu, vert, rouge, proche infra-rouge*) sont utilisées comme attributs et paraissent être suffisantes pour se focaliser sur les régions de bâti les plus probables (Figure 1b).

3.2 Détection de contours

L'extraction descendante de motifs rectangulaires est guidée par l'information de gradients dans l'image (Section 4). À cet effet, une carte de gradient améliorée G est calculée, en combinant une image de gradients avec un détecteur de segments, à savoir LSD [15]. Nous supposons que les bâtiments ont de faibles courbures et qu'ils peuvent être représentés par des rectangles : leurs bords peuvent être décomposés en segments. L'algorithme LSD permet d'avoir des segments non exhaustifs mais robustes et fiables, avec des temps de calcul réduits et une précision sub-pixellaire. La valeur de gradient améliorée correspond à la valeur moyenne du module de gradient (rééchelonnée entre 0 et 1) et la valeur LSD binaire. La figure 1c montre que les bords de bâtiments sont correctement détectés et avec une grande précision. Cette approche peut conduire toutefois à de faux positifs situés à l'intérieur des toits, de la végétation, des zones d'ombre etc. Cependant, ils peuvent être éliminés par l'étape suivante d'extraction de rectangles (Section 4.1).

3.3 Partitionnement de l'image

L'extraction descendante de rectangles nécessite l'exploration exhaustive de l'espace de configurations de rectangles afin de trouver celles qui correspondent le mieux aux informations de bas niveau extraites de l'image. Le processus peut être accéléré en limitant l'espace de recherche aux zones les plus pertinentes de l'image [16, 17], ici les zones de bâti. Ainsi, la carte de probabilité de la classe *bâtiment* (Section 3.1) est utilisée. Elle est lissée et binarisée en utilisant une segmentation de l'image satellitaire originale [18]. L'image $\mathcal{I}$ est donc segmentée en n régions R_i, et ensuite partagée en deux sous-ensembles disjoints : $\mathcal{I} = \mathcal{S}_b \cup \mathcal{S}_{nb}$, avec $\mathcal{S}_b$ et $\mathcal{S}_{nb}$ l'ensemble des pixels *bâtiment* et *non-bâtiment*, de tailles n_b et $n-n_b$ respectivement (zones orange et violette sur la Figure 1d). On a :

$$\begin{cases} \mathcal{S}_b & = \bigcup_{i=1..n_b} R_i \ / \ \overline{p_{R_i}} > \tau \\ \mathcal{S}_{nb} & = \bigcup_{j=1..n-n_b} R_j \ / \ \overline{p_{R_j}} \leq \tau \end{cases} \quad (1)$$

$\overline{p_{R_i}}$ est la probabilité moyenne de la classe *bâtiment* pour R_i et τ le seuil de probabilité à partir duquel une région sera utilisée pour l'extraction de rectangles. En pratique, τ est choisi suffisamment bas (1) pour être sélectionné de façon automatique et (2) pour éviter d'éliminer des régions potentiellement candidates. Par conséquent, τ est fixé de façon à éliminer les zones de végétation et les régions du sol présentant des réponses spectrales similaires. Pour cela, on utilise l'indice de végétation de différence normalisée (NDVI). Un sous-ensemble $\mathcal{S}_v$ de 100 pixels de végétation est sélectionné aléatoirement parmi les 1000 pixels présentant les plus fortes valeurs de NDVI. On a alors : $\tau = \overline{p_{\mathcal{S}_v}}$. La figure 1d montre une carte du partitionnement binaire de l'image satisfaisante. Les zones de bâti sont conservées avec une focalisation grossière, réduisant considérablement les temps de calcul (jusqu'à 40% de l'image peut être supprimée).

4 Détection descendante

Nous supposons que chaque empreinte de bâtiment peut être approchée par un rectangle ou l'union de plusieurs rectangles qui se chevauchent légèrement. Nous présentons d'abord l'extraction des rectangles et ensuite la segmentation finale des bâtiments.

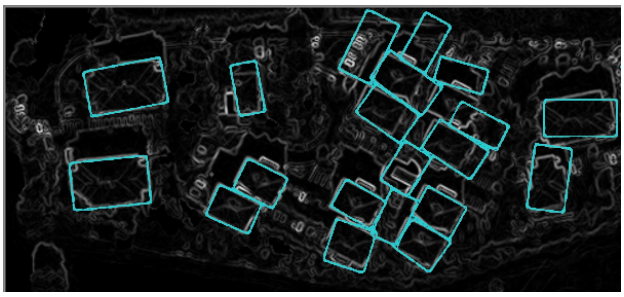
4.1 Extraction de rectangles

Le cadre stochastique des Processus Ponctuels Marqués (PPM) est utilisé pour modéliser l'image par un ensemble de rectangles [19]. Les PPM ont en effet montré toute leur pertinence et leur compétitivité pour l'extraction d'objets paramétriques dans de nombreux problèmes de vision et en particulier pour l'extraction de bâtiments sur des images géospatiales [20]. Les processus ponctuels markoviens sont des modèles probabilistes introduits par [21] afin d'étendre les champs de Markov traditionnels grâce à un formalisme orienté-objet. Ces outils mathématiques exploitent des variables aléatoires dont les réalisations sont des configurations d'objets paramétriques, chaque objet étant assigné à un point positionné dans la scène [17]. Le nombre d'objets est lui-même une variable aléatoire. L'une des forces de ce formalisme est la possibilité d'introduire des interactions spatiales plus ou moins complexes entre les objets tout en imposant des contraintes de régularisation globale sur la scène d'intérêt. Un processus ponctuel est défini par des objets paramétriques (ici des rectangles qui créent une configuration x), une énergie (qui calcule la cohérence entre la configuration x et l'image observée, et prend en compte les interactions entre les rect-

angles $r \in x$) et un échantillonneur (qui trouve la configuration minimisant l'énergie). Cette minimisation est complexe puisque le nombre d'objets dans la configuration est inconnu, et que l'énergie n'est pas convexe. Un échantillonneur stochastique est utilisé pour trouver le minimum global de U [19]. En particulier, il s'agit de l'algorithme de Métropolis-Hastings-Green [22] qui appartient à la famille des méthodes de Monte-Carlo par Chaîne de Markov (MCMC). L'énergie U mesure la qualité d'une configuration x et est composée d'un terme d'attache aux données U_d et d'un terme de régularisation (a priori) U_p telle que :

$$U(x) = \sum_{r \in x} U_d(r) + \beta \sum_{\substack{r_i, r_j \in x \\ i \neq j}} U_p(r_i, r_j)\,, \qquad (2)$$

avec $\beta \in \mathbf{R}^+$, le coefficient de pondération.



(a)

(b)

FIGURE 2 – Extraction de rectangles par Processus Ponctuels Marqués. (a) Pour un paramétrage donné (cyan – τ_G=60, C=2300), en superposition sur la carte de gradient améliorée G. Des faux positifs apparaîssent sur les routes puisque leurs bords correspondent aux bords des bâtiments de part et d'autre de la route ou bien sur des files de voitures puisqu'elles renvoient des forts gradients sur les bords. (b) Quand le modèle est paramétré pour des bâtiments larges (fortes valeurs de γ), un fort taux de fausses détections existe et de grands rectangles couvrent des ensembles de bâtiments.

Attache aux données. Il est basé sur le fait que les bords d'un rectangle doivent correspondre à des pixels de forte valeur de gradient. Par conséquent, nous utilisons la carte de gradient améliorée G, sans formuler aucune hypothèse sur la couleur des toits ou l'ombre du bâtiment, afin de conserver la généricité de l'approche. L'énergie d'attache aux données est définie par :

$$U_d(r) = C - U_G(r). \qquad (3)$$

U_G est le terme d'énergie basé sur le gradient, et C un paramètre de pénalité pour l'ajout de nouveaux rectangles (on favorise les configurations avec un nombre minimal de rectangles). C est une constante pour une zone donnée et est fixée manuellement en fonction des valeurs de U_G.

$$U_G(r) = \sum_{i=1}^{4} \frac{l_{\tau_G}(r_i)}{l(r_i)} \times \prod_{i=1}^{4} G(r_i). \qquad (4)$$

r_i est un côté du rectangle r, $l(r_i)$ sa longueur, et $G(r_i)$ sa valeur moyenne de gradient. $l_{\tau_G}(r_i)$ est la longueur maximale de r_i où les valeurs du gradient sont supérieures à un seuil donné $\tau_G \in [50, 100]$.
Le premier terme de l'équation assure que les bords du rectangle ont continuellement de fortes valeurs de gradient, tandis que le second vise à favoriser les valeurs les plus élevées de G sur tous les bords du rectangle. Cette formulation permet de ne pas être pénalisé par des occultations par la végétation et les zones d'ombre. Une faible valeur de τ_G est suffisante pour se concentrer sur les empreintes de bâtiments. En pratique, une évaluation de deux bords orthogonaux des rectangles doit être suffisante pour une bonne détection approprié, mais cela pourrait conduire à de fausses détections sur les marques routières par exemple.

Régularisation. L'énergie U_p est utilisée pour favoriser des configurations souhaitées entre rectangles voisins et pénaliser les autres. Les objets sont considérés comme voisins s'ils s'intersectent. Un seul bâtiment peut être représenté par un seul rectangle (de voisinage vide) tandis que pour des blocs de bâtiments, l'arrangement des rectangles nécessaires pour les représenter se compose de rectangles contigus ou qui se chevauchent légèrement. Pour obtenir les configurations souhaitées, l'énergie a un seul terme de répulsion. Pour une paire d'objets voisins r_i et r_j, elle correspond tout simplement à leur surface d'intersection S_{inter} [20, 3], ce qui évite une redondance de détection pour un même bâtiment :

$$U_p(r_i, r_j) = S_{inter}(r_i, r_j), \qquad (5)$$

Optimisation. La recherche de la configuration de rectangles qui minimise l'énergie non convexe U dans un espace de dimension variable se fait à l'aide d'un échantillonneur par saut réversible de Monte-Carlo par Chaîne de Markov (RJMCMC) couplé à une relaxation stochastique. Un des avantages de cet algorithme itératif est qu'il ne dépend pas de l'état initial. Plus de détails peuvent être trouvés dans [19]. Les performances en temps de calcul et en stabilité de la convergence (dans le cas de grandes

scènes) peuvent être limitées, c'est pourquoi nous utilisons le partitionnement grossier de l'image pour interdire la création de rectangles dans les zones *Non-bâtiment*. On fixe également un a priori sur les tailles maximales des bâtiments pour limiter l'exploration de l'espace des rectangles de l'image.

FIGURE 3 – Multiples détections par PPM sur la même zone d'intérêt (une couleur par réalisation) : des fausses détections persistent mais sont compensées par la fusion des différents résultats de détection de rectangles (Figure 1e).

Carte de probabilité de bâtiments. Le modèle proposé est efficace pour la détection des parties principales du bâtiment, mais souffre de deux limitations principales : la simplicité du modèle (de fausses détections existent) et la sensibilité au paramétrage (inadaptation quand les objets recherchés sont de taille variable). Puisque le résultat d'une unique optimisation ne fournit qu'une carte binaire de bâtiments, on ne peut l'utiliser directement comme carte de probabilité de bâtiments pour la classification finale. Pour lever ces deux limitations, différents paramétrages sont testés et leurs résultats sont fusionnés. Les intervalles de variation des différents paramètres sont fixés de façon expérimentale. Ensuite, les intervalles sont échantillonnés de façon uniforme selon l'impact de chaque paramètre en deux, trois ou quatre sous-intervalles pour respectivement τ_G, C et γ (le paramètre d'élongation du rectangle, cf. [3]). Ceci permet d'obtenir un ensemble de 24 configurations distinctes, noté $\mathcal{R}$. Enfin, chaque pixel i de l'image $\mathcal{I}$ est associé à une probabilité P_{TD} calculée à partir des rectangles r contenant ce pixel.

$$\forall i \in \mathcal{I}, \quad P_D(i) = \frac{1}{card\,\mathcal{R}} \sum_{r \in \mathcal{R}} 1_{\{i \in r\}}, \tag{6}$$

où $1_{\{.\}}$ désigne la fonction indicatrice. La figure 3 montre des résultats sur une zone résidentielle. On observe que l'extraction de rectangles n'est pas parfaite en termes de nombre d'objets et de précision de localisation. Toutefois, malgré la simplicité du modèle, on observe que la plupart des bâtiments sont détectés et que le processus n'est pas perturbé par les zones d'ombre. La carte de probabilité donnée par l'équation 6 est illustrée sur la figure 1e. On peut voir que l'accumulation des multiples extractions de rectangles fournit une carte fiable de probabilité de bâtiments. Les faibles valeurs de confiance correspondent à des fausses détections ou à des bords de bâtiments, alors que les valeurs les plus élevées correspondent aux toits de bâtiments.

4.2 Segmentation finale

La décision finale est prise au niveau pixellaire et de façon non-supervisée par minimisation d'énergie. Cette énergie globale d'une coupure $\mathcal{C}$ est composée d'un terme d'attache aux données E_d, dérivé de l'approche de classification ascendante, et une énergie de régularisation notée E_p et définie par la carte de probabilité obtenue par l'approche descendante. On a alors :

$$E(C) = \sum_{p \in \mathcal{I}} E_{\mathrm{d}}(p) + \lambda \sum_{p,q \in \mathcal{I}} E_{\mathrm{p}}(p,q), \tag{7}$$

avec p, q des pixels de l'image $\mathcal{I}$, $\lambda > 0$ un paramètre de pondération entre les deux termes d'énergie ascendante et descendante et (p, q) une paire de pixels voisins.
L'énergie est minimisée par une coupure de graphe [23, 24] qui renvoie une solution exacte. Le terme d'attache aux données E_d qui est défini comme suit :

$$E_{\mathrm{d}}(p) = \log(1 + k \times P_{\text{bât}}(p)), \tag{8}$$

avec $P_{\text{bât}}$ la probabilité (ascendante, par SVM) d'appartenance à la classe *bâtiment*, et k un paramètre de confiance dans le processus de labellisation (ici k=1000).
E_{p} vise à favoriser des pixels voisins ayant des probabilités *bâtiment* similaires calculées par l'approche descendante (P_{D}), et défavorise fortement les paires ayant des valeurs très différentes. Il est défini comme suit :

$$E_{\mathrm{p}}(p,q) = \exp\left(-\frac{\max(P_D) - |P_{\mathrm{D}}(p) - P_{\mathrm{D}}(q)|}{\sigma}\right), \tag{9}$$

avec $\max(P_D)$ la probabilité descendante maximale, σ le facteur de normalisation permettant d'avoir des valeurs E_p similaires quelque soit la performance de l'étape d'extraction multiple de rectangles.

5 Résultats et discussions

5.1 Données

Notre méthode a été évaluée sur une zone mixte urbaine-rurale au sud de Toulouse. Elle couvre 1,5 km^2 et présente un terrain en pente. Une grande variabilité de bâtiments est présente : de larges bâtiments (une école, un hôpital), 230 petites maisons individuelles, avec des matériaux de toits parfois très différents [5]. Une zone de végétation est également présente avec de la forêt et des terrains cultivés. Une image Pléiades multispectrale orthorectifiée a été utilisée.

5.2 Résultats

Pour une analyse détaillée, 3 sous-zones d'intérêt ont été sélectionnées. Le temps de calcul pour chaque zone est d'environ 50 minutes et est principalement dû à l'extraction multiple de rectangles par PPM. Chaque paramétrage de PPM correspond à 15 millions d'itérations et nécessite

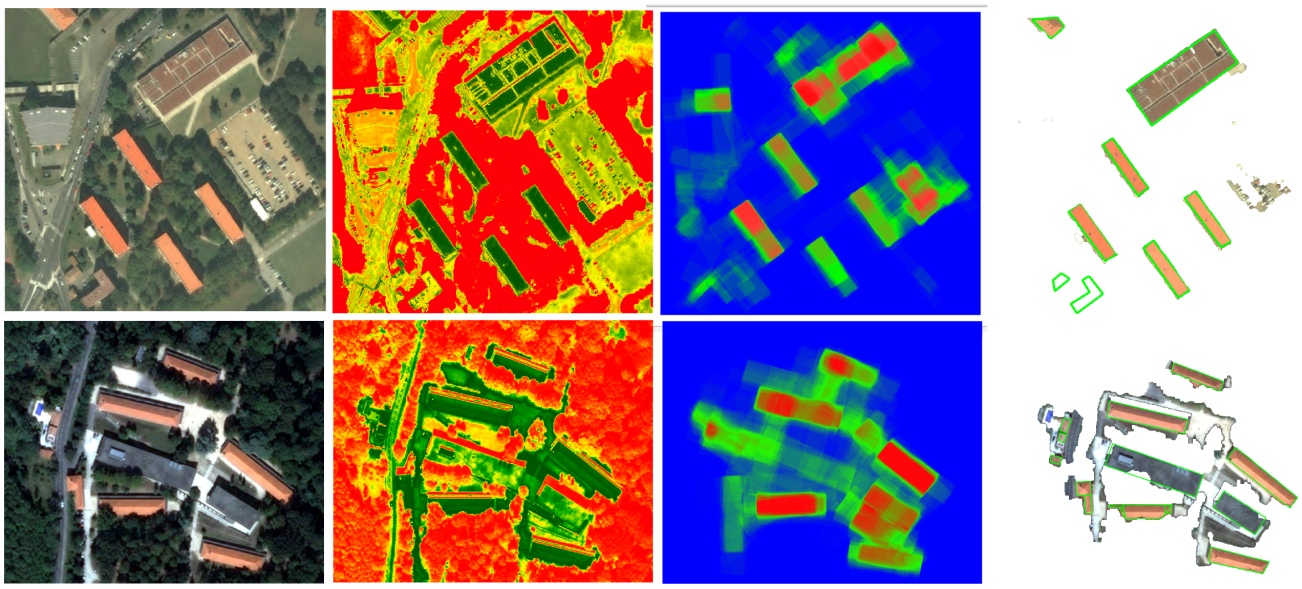


FIGURE 4 – Résultats sur des zones d'intérêt. **De haut en bas** : *Zone 2* et *Zone 3*. **De gauche à droite** : Image satellitaire – carte de probabilité par classification ascendante – carte de probabilité par PPM multiples – Segmentation finale, comparée à la Vérité Terrain (voir Figure 1 pour la légende).

approximativement 2 minutes. La figure 4 montre des résultats de détection de bâtiments sur les deux sous-zones non présentées sur la Figure 1.
Puisque la classification SVM a été réalisée uniquement à partir des quatre bandes spectrales, on peut voir que la classification ascendante est très sensible aux matériaux et à l'orientation des toits conduisant à différentes probabilités pour un même toit. De plus, certaines confusions apparaissent entre les toits et les routes comme ils ont des réponses spectrales similaires. Par contre, on peut également voir que la carte de probabilité par PPM est robuste aux zones d'ombre et met bien en évidence les bâtiments même si leur localisation reste grossière (*Zones 1* et *2*). La plupart des bâtiments est détectée, même ceux se trouvant dans les zones boisées. Cependant, certains faux positifs demeurent. Ils correspondent aux zones non bâties où les probabilités ascendante et descendante des *bâtiments* sont fortes (*Zone 3*) : il s'agit des zones au sol avec un contraste élevé telles que les piscines, les files de véhicules dans les parkings ou avec une réflectance très discriminante (*Zones 1*, *2* et *3*, respectivement). Toutefois, la combinaison des approches permet de minimiser les erreurs individuelles et l'utilisation de modèles plus avancés en particulier pour la classification devrait résoudre ces problèmes.
L'approche délimite avec précision la plupart des bâtiments. Une analyse quantitative a été effectuée pour évaluer la qualité de l'approche proposée [25]. La classification binaire finale a été comparée à une carte de vérité terrain extraite de la BDTOPO© IGN pouvant être parfois imprécise. Des taux de rappel et de précision ont été calculés via une évaluation par pixel (Tableau 1). Ils confirment les impressions visuelles décrites ci-dessus.

Zone	**Rappel** (%)	**Précision** (%)
1	89	81,5
2	90,5	80,5
3	94	67

TABLE 1 – Qualité de l'extraction de bâtiments par notre méthode, comparée à une base de données topographique de référence.

6 Conclusion

Nous avons proposé une chaîne de traitements simple et efficace pour combiner des approches ascendante et descendante pour la détection de bâtiments à partir d'une image optique THR monoculaire. Pour chaque étape, des méthodes de l'état-de-l'art ont été adoptées. L'intérêt de notre méthode est de nécessiter peu de paramètres et peu de connaissances a priori ce qui la rend facilement applicable à grande échelle. En outre, grâce à la fusion des résultats de multiples processus ponctuels marqués, on s'affranchit du problème délicat du paramétrage des modèles stochastiques et c'est là la contribution principale de cet article. Enfin, ce travail est fondé sur des logiciels libres [26, 27] ce qui rend le travail facilement reproductible.
Les résultats sont satisfaisants tant en termes de détection que de délimitation des bâtiments. D'autres tests de la méthode devront être réalisés sur différents types de paysage et d'autres capteurs.
La méthode est actuellement en cours de comparaison avec les approches développées par [1, 13, 28] afin de mieux quantifier l'apport des différentes techniques employées et en particulier de la pertinence de la combinaison entre ascendantes et descendantes.

Références

[1] A. O. Ok, "Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 86, pp. 21–40, 2013.

[2] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny, "Structural approach for building reconstruction from a single DSM," *PAMI*, vol. 32, no. 1, pp. 135–147, 2010.

[3] M. Brédif, O. Tournaire, B. Vallet, and N. Champion, "Extracting polygonal building footprints from Digital Surface Models : A fully-automatic global optimization framework," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 77, pp. 57–65, 2013.

[4] J. Wegner, R. Hänsch, A. Thiele, and U. Soergel, "Building detection from one orthophoto and high-resolution insar data using conditional random fields," *JSTARS*, vol. 4, no. 1, pp. 83–91, 2011.

[5] A. Le Bris and N. Chehata, "Change detection in a topographic building database using submetric satellite images," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII-3/W22, pp. 25–30, 2011.

[6] C. Senaras, M. Ozay, and F. Yarman Vural, "Building detection with decision fusion," *JSTARS*, vol. 6, no. 3, pp. 1295–1304, 2013.

[7] K. Karantzalos and N. Paragios, "Recognition-driven two-dimensional competing priors toward automatic and accurate building detection," *TGRS*, vol. 47, no. 1, pp. 133–144, 2009.

[8] C. Benedek, X. Descombes, and J. Zerubia, "Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics," *PAMI*, vol. 34, no. 1, pp. 33–50, 2012.

[9] D. Grosgeorge, C. Petitjean, and S. Ruan, "Segmentation d'images par coupe de graphe avec a priori de forme," in *RFIA*, 2012.

[10] H. Harzallah, F. Jurie, and C. Schmid, "Combining efficient object localization and image classification," in *ICCV*, 2009.

[11] L. Ladicky, P. Sturgess, K. Alahari, C. Russell, and P. Torr, "What, Where & How Many ? Combining object detectors and CRFs," in *ECCV*, 2010.

[12] M. Maire, S. Yu, and P. Perona, "Object detection and segmentation from joint embedding of parts and pixels," in *ICCV*, 2011.

[13] D. Chai, W. Förstner, and M. Ying Yang, "Combine Markov Random Fields and Marked Point Processes to extract building from remotely sensed images," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I-3, pp. 365–370, 2012.

[14] T.-F. Wu, C.-J. Lin, and R. Weng, "Probability estimates for multi-class classification by pairwise coupling," *JMLR*, vol. 5, pp. 975–1005, 2004.

[15] R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD : A fast line segment detector with a false detection control," *PAMI*, vol. 32, no. 4, pp. 722–732, 2010.

[16] Z. Tu and S. Zhu, "Image segmentation by Data-Driven Markov Chain Monte Carlo," *PAMI*, vol. 24, no. 5, pp. 657–673, 2002.

[17] Y. Verdié and F. Lafarge, "Detecting parametric objects in large scenes by Monte Carlo sampling," *IJCV*, vol. 106, no. 1, pp. 57–75, 2014.

[18] L. Guigues, J. Cocquerez, and H. L. Men, "Scale-sets image analysis," *IJCV*, vol. 68, no. 3, pp. 289–317, 2006.

[19] X. Descombes, Ed., *Applications de la géométrie stochastique à l'analyse d'images*. Lavoisier, 2011.

[20] M. Ortner, X. Descombes, and J. Zerubia, "Building outline extraction from digital elevation models using marked point processes," *IJCV*, vol. 72, no. 2, pp. 107–132, 2007.

[21] A. Baddeley and M. V. Lieshout, "Stochastic geometry models in high-level vision," *Statistics and Images*, vol. 1, no. 2, pp. 233–258, 1993.

[22] P. Green, "Reversible Jump Markov Chain Monte-Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[23] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," in *ICCV*, 1999, pp. 377–384.

[24] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts," *PAMI*, vol. 26, no. 2, pp. 147–159, 2004.

[25] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *JSTARS*, vol. 2, no. 1, pp. 11–20, 2009.

[26] M. Brédif and O. Tournaire, "Librjmcmc : An open-source generic c++ library for stochastic optimization," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIX-B3, pp. 259–264, 2012. [Online]. Available : http ://librjmcmc.ign.fr

[27] E. Christophe and J. Inglada, "Open source remote sensing : Increasing the usability of cutting-edge algorithms," *IEEE Geoscience and Remote Sensing Newsletter*, vol. 35, no. 5, pp. 9–15, 2009.

[28] J. Niemeyer, F. Rottensteiner, and U. Soergel, "Contextual classification of lidar data and building object detection in urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 152–165, 2013.

Bibliographie

P. Andrieux, A. Hatier, J. Asseline, G. de Noni, and M. Voltz. Predicting infiltration rates by classifying soil surface features in a mediterranean wine-growing area. In *International Symposium 'The Significance of Soil Surface Characteristics in Soil Erosion. COST 623 Soil Erosion and Global Change workshop*, Strasbourg, 2001. (Cité page 37.)

P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5) :898–916, may 2011. ISSN 0162-8828. (Cité page 66.)

P. Arbeláez, B. Hariharan, C. Gu, S. Gupta, L. Bourdev, and J. Malik. Semantic segmentation using regions and parts. In *CVPR*, 2012. (Cité page 66.)

M. Baatz and M. Schape. *Multiresolution segmentation - An optimization approach for high quality multi-scale image segmentation*. Strobl, J., Blaschke, T., Griesebner, G. (Eds) Angewandte Geographische Informations - Verarbeitung XII. Wichmann Verlag, Karlsruhe, 2000. (Cité page 69.)

R. Battiti. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, 5(4), jul 1994. (Cité page 60.)

B. Beguet, N. Chehata, S. Boukir, and D. Guyon. Retrieving forest structure variables from very high resolution satellite images using an automatic method. In *ISPRS International Society for Photogrammetry and Remote Sensing*, volume I (7), pages 1–6, Melbourne, 2012. (Cité page 63.)

Benoit Beguet. *Caractérisation et cartographie de la structure forestiére á partir dâimages satellitaires á trés haute résolution spatiale.* PhD thesis, Ecole doctorale Montaigne Humanités- Université Michel de Montaigne, 2014. (Cité pages 40, 56 et 62.)

Benoit Beguet, Dominique Guyon, Samia Boukir, and Nesrine Chehata. Automated retrieval of forest structure variables based on multi-scale texture analysis of VHR satellite imagery. *ISPRS International Journal of Photogrammetry and Remote Sensing*, 96(0) :164–178, 2014. ISSN 0924-2716. (Cité page 63.)

C. Benedek, X. Descombes, and J. Zerubia. Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. *PAMI*, 34(1) :33–50, 2012. (Cité page 66.)

U.C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. Multiresolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4) : 239–258, 2004. (Cité page 69.)

T. Blaschke. Object based image analysis for remote sensing. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 65(1) :2–16, 2010. ISSN 0924-2716. (Cité page 66.)

S. Bo, L. Ding, H. Li, F. Di, and C. Zhu. Mean shiftâbased clustering analysis of multispectral remote sensing imagery. *International Journal of Remote Sensing*, 30(4) :817–827, 2009. (Cité page 71.)

S. Boukir, L. Guo, and N. Chehata. Classification of remote sensing data using margin-based ensemble methods. In *IEEE International Conference on Image Processing ICIP 2013*, pages 2602–2606, Melbourne, Australia, Sep 2013. (Cité pages 81 et 83.)

Y. Boykov and D. Huttenlocker. A New Bayesian Framework for Object Recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 517–523, 1999. (Cité page 70.)

Y. Boykov and M-P. Jolly. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D images. In *IEEE Conference on Computer Vision*, volume 1, pages 105–112, jul 2001. (Cité page 70.)

Y. Boykov, O. Veksler, and R. Zabih. Markov random fields with efficient approximations. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 648–655, 1998. (Cité page 70.)

Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. In *ICCV*, pages 377–384, 1999. (Cité page 70.)

L. Breiman. Random forests. *Machine Learning*, 45(1) :5–32, 2001. (Cité pages 61, 78 et 79.)

F. Bretar and N. Chehata. Génération de modéles numériques de terrain par fusion de données lidar et image. *Revue Traitement du Signal*, 26(2), 2009. (Cité page 44.)

F. Bretar and N. Chehata. Terrain modelling from lidar range data in natural landscapes : a predictive and bayesian framework. *IIEEE Transactions on Geoscience and Remote Sensing*, 48(3) :1568–1578, 2010. (Cité pages 42 et 87.)

Xavier Briottet, Nesrine Chehata, Rosa Oltra-Carrio, Arnaud Le Bris, and Christiane Weber. *Application de l'imagerie optique aux milieux urbains*. Télédétection pour l'Observation des Surfaces Continentales, Tome 4 : Applications thématiques : Environnement et Risques. ISTE Editeur, Elsevier, Système Terre â Environnement, 2016. (Cité page 34.)

L. Bruzzone and S. B. Serpico. A technique for feature selection in multiclass problem. *International Journal of Remote Sensing*, 21(3) :549–563, 2000. (Cité page 60.)

L. Bruzzone, M. Chi, and M. Marconcini. A novel transductive svm for semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(11) :3363–3373, Nov 2006. ISSN 0196-2892. (Cité page 82.)

C. Burnett and T. Blaschke. a multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling*, 168(3) : 233–249, 2003. (Cité page 69.)

G. Camps-Valls and L. Bruzzone. Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43 (6) :1351–1362, jun 2005. (Cité page 55.)

S. Cang and Y. Hongnian. Mutual information based input feature selection for classification problems. *Decision Support Systems*, 54 :691–698, 2012. (Cité page 60.)

D. Chai, W. Förstner, and M. Ying Yang. Combine markov random fields and marked point processes to extract building from remotely sensed images. *ISPRS Annals*, I-3 :365–370, 2012. (Cité page 67.)

Chein-I Chang and Su Wang. Constrained band selection for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6) :1575–1585, jun 2006. (Cité page 60.)

Q. Chang, C.-I.and Du, T.-L. Sun, and M. L. G. Althouse. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (6) :2631–2641, nov 1999. (Cité page 60.)

Y.-L. Chang, K.-S. Chen, B. Huang, W.-Y. Chang, J.A. Benediktsson, and L. Chang. A parallel simulated annealing approach to band selection for high-dimensional remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(3) :579–590, sep 2011. (Cité page 62.)

Olivier Chapelle, Bernhard Schlkopf, and Alexander Zien. *Semi-Supervised Learning*. The MIT Press, 1st edition, 2010. ISBN 0262514125, 9780262514125. (Cité page 95.)

N. Chehata. *Modélisation 3D de scènes urbaines à partir d'images satellitaires à très haute résolution*. PhD thesis, Université Paris Descartes, France, Jun 2005. (Cité page 74.)

N. Chehata and F. Bretar. Terrain modeling from lidar data : Hierarchical K-means filtering and markovian regularization. In *IEEE International Conference on Image Processing (ICIP)*, San Diego, 12-15 Octobre 2008. (Cité pages 73 et 87.)

N. Chehata, M. Pierrot Deseilligny, F. Jung, and G. Stamon. Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings. In *MVA*, pages 636–640, 2002. (Cité page 74.)

N. Chehata, F. Jung, M. Pierrot Deseilligny, and G. Stamon. A region-based matching approach for 3D-roof reconstruction from HR satellite stereo pairs. In *DICTA*, pages 889–898, 2003. (Cité page 75.)

N. Chehata, N. David, and F. Bretar. Lidar data classification using hierarchical K-means clustering. In *International Archives of Photogrammetry, Remote*

Sensing and Spatial Information Sciences, volume 37 (Part 3B), pages 325–330, Pekin, Juil 2008. (Cité pages 42 et 87.)

N. Chehata, L. Guo, and C. Mallet. Contribution of airborne full-waveform lidar and image data for urban scene classification. In *IEEE International Conference on Image Processing (ICIP)*, Caire, 7-10 Novembre 2009a. (Cité page 63.)

N. Chehata, L. Guo, and C. Mallet. Airborne lidar feature selection for urban classification using random forests. In *Laserscanning 2009, IAPRS*, volume 38 (Part 3/W8), pages 207–212, Paris, sep 2009b. (Cité pages 42, 54, 58 et 63.)

N. Chehata, A. Le Bris, and P. Lagacherie. Comparison of VHR panchromatic texture features for tillage mapping. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2013*, pages 3128–3131, 2013. (Cité pages 56 et 65.)

N. Chehata, A. Le Bris, and S. Najjar. Contribution of band selection and fusion for hyperspectral classification. In *IEEE WHISPERS*, Lausanne, Suisse, Jun 2014a. (Cité page 64.)

N. Chehata, C. Orny, S.Boukir, and D. Guyon. Object-based change detection in wind-storm damaged forest using high resolution multispectral images. *IInternational Journal of Remote Sensing (IJRS)*, 35(13) :4758–4777, 2014b. (Cité pages 40, 65, 71, 77 et 86.)

N. Chehata, K. Ghariani, A. Le Bris, and P. Lagacherie. Délimitation des parcelles agricoles par classification d'images pléiades. *Revue Française de Photogrammétrie et de Télédétection*, 209 :165–171, 2015. (Cité pages 56, 65, 85 et 102.)

D. Comaniciu, V. Ramesh, and P. Meer. The variable bandwidth mean shift and data-driven scale selection. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 1, pages 438–445, 2001. (Cité page 72.)

Dorin Comaniciu, Peter Meer, and Senior Member. Mean shift : A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 :603–619, 2002. (Cité pages 71, 72 et 86.)

N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In IEEE Computer Society, editor, *International Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages pp 886–893, San Diego, United States, 2005. (Cité page 57.)

M. Dalponte, L. Bruzzone, and D. Gianelle. Fusion of hyperspectral and lidar remote sensing data for classification of complex forest areas. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5) :1416–1427, May 2008. ISSN 0196-2892. (Cité page 54.)

M. Dalponte, H.O. Orka, and L.T. Ene. Tree crown delineation and tree species classification in boreal forests using hyperspectral and als data. *Remote Sensing of Environment*, 140 :306–317, 2014. (Cité page 44.)

Steve De Backer, Pieter Kempeneers, Walter Debruyn, and Paul Scheunders. Band selection for hyperspectral remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1–4, 2005. (Cité page 62.)

Carole Delenne, Sylvie Durrieu, Gilles Rabatel, and Michel Deshayes. From pixel to vine parcel : A complete methodology for vineyard delineation and characterization using remote-sensing data. *Computers and Electronics in Agriculture*, 70(1) :78–83, 2010. ISSN 0168-1699. (Cité page 67.)

J. Demantke, C. Mallet, N. David, and B. Vallet. Â dimensionality-based scale selection in 3d lidar point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(5/W12) :97–102, 2011. (Cité page 57.)

B. Demir, L. Minello, and L. Bruzzone. An effective strategy to reduce the labeling cost in the definition of training sets by active learning. *IEEE Geoscience and Remote Sensing Letters*, 11(1) :79–83, Jan 2014. (Cité pages 82 et 94.)

B.J. Devereux, G.S. Amable, and C. Costa Posada. An efficient image segmentation algorithm for landscape analysis. *International Journal of Applied Earth Observation and Geoinformation*, 6(1) :47–61, 2004. ISSN 0303-2434. (Cité pages 68 et 69.)

R. Díaz-Uriarte and S. A. De Andres. Gene selection and classification of microarray data using Random Forest. *BMC bioinformatics*, 7(3) :1–13, 2006. (Cité page 60.)

R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973. pp. 98â105. (Cité page 54.)

Pablo A. Estévez, Michel Tesmer, Claudio A. Perez, and Jacek M. Zurada. Normalized mutual information feature selection. *IEEE Transactions on Neural Networks*, 20(2) :189–201, feb 2009. (Cité page 62.)

M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton. Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101(3) :652–675, March 2013. ISSN 0018-9219. (Cité pages 67 et 78.)

M. Fauvel, A. Zullo, and F. Ferraty. Nonlinear parsimonious feature selection for the classification of hyperspectral images. In *Proc. of the 6th Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS'14)*, jun 2014. (Cité page 60.)

P.F. Felzenszwalb and D.P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59 :2004, 2004. (Cité page 70.)

A. Ferraz., C. Mallet, and N. Chehata. Détection de la desserte forestière en milieu montagneux par lidar aéroporté. *Revue Française de Photogrammétrie et de Télédétection*, 211-212 :103–117, 2015. (Cité page 88.)

A. Ferraz, C. Mallet, and N. Chehata. Large-scale road detection in forested mountainous areas using airborne topographic lidar data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 112 :23–36, 2016. ISSN 0924-2716. (Cité pages 42 et 88.)

António Ferraz, Frédéric Bretar, Stéphane Jacquemoud, Gil Gonçalves, Luisa Pereira, Margarida Tomé, and Paula Soares. 3-D mapping of a multi-layered

Mediterranean forest using {ALS} data. *Remote Sensing of Environment*, 121 : 210–223, 2012. ISSN 0034-4257. (Cité page 73.)

S. Filin and N. Pfeifer. Neighborhood systems for airborne laser data. *Photogrammetric Engineering & Remote Sensing*, 71(6) :743–755, 2005. (Cité page 57.)

A.O. Finley, S. Banerjee, and B.D Cook et al. Hierarchical bayesian spatial models for predicting multiple forest variables using waveform lidar, hyperspectral imagery, and large inventory datasets. *International Journal of Applied Earth Observation and Geoinformation*, 22 :147–160, 2013. (Cité page 44.)

R.A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2) :179–188, 1936. ISSN 2050-1439. (Cité page 60.)

G.M. Foody and A. Mathur. Toward intelligent training of supervised image classifications : directing training data acquisition for SVM classification. *Remote Sensing of Environment*, 93(1â2) :107–117, 2004. (Cité page 80.)

Z. L. Frogbrook and M. A. Oliver. Identifying management zones in agricultural fields using spatially constrained classification of soil and ancillary data. *Soil Use and Management*, 23(1) :40–51, 2007. (Cité page 67.)

A. Fu and H. Yan. A new probabilistic relaxation method based on probability space partition. In *Pattern Recognition*, volume 30(11). Elsevier Science Ltd, 1997. (Cité page 70.)

K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 1 :32–40, 1975. (Cité page 71.)

S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 :721–741, 1984. (Cité page 70.)

R. Genuer, J.-M. Poggi, and C. Tuleau-Malot. Variable selection using Random Forests. *Pattern Recognition Letters*, 31 :2225–2236, 2010. (Cité page 61.)

B. Georgescu, I. Shimshoni, and P. Meer. Mean shift based clustering in high dimensions : A texture classification example. *IEEE International Conference on Computer Vision*, 1 :456–463, 2003. (Cité page 72.)

R.G. Von Gioi, j. Jakubowicz, J.M. Morel, and G. Randall. LSD : A fast line segment detector with a false detection control. *PAMI*, 32(4) :722–732, 2010. (Cité pages 66, 76 et 77.)

A. Golovinskiy, V.G. Kim, and T. Funkhouser. Shape-based recognition of 3D point clouds in urban environments. *International Conference on Computer Vision (ICCV)*, sep 2009. (Cité page 73.)

R.C. Gonzales and R.E. Woods. *Digital image processing*. Pearson Education Inc., 2008. (Cité page 67.)

S. Gould, T. Gao, and D. Koller. Region-based segmentation and object detection. In *NIPS*, 2009. (Cité page 66.)

S.E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on gabor filters. *IEEE Transactions on Image Processing*, 11(10) :pp 1160–1167, 2002. (Cité page 57.)

Laurent Guigues, Jean Pierre Cocquerez, and Hervé Le Men. Scale-Sets Image Analysis. *International Journal of Computer Vision*, 68(3) :289–317, 2006. ISSN 1573-1405. (Cité page 69.)

Baofeng Guo, R.I. Damper, Steve R. Gunn, and J.D.B. Nelson. A fast separability-based feature-selection method for high-dimensional remotely sensed image classification. *Pattern Recognition*, 41 :1653–1662, 2008. (Cité page 60.)

L. Guo, S. Boukir, and N. Chehata. Support vectors selection for supervised learning using an ensemble approach. In *IAPR International Conference on Pattern Recognition (ICPR)*, pages 37–40, Istanbul, 23-26 Août 2010a. (Cité page 81.)

L. Guo, N. Chehata, and S. Boukir. A two-pass random forests classification of airborne lidar and image data on urban scenes. In *IEEE International Conference on Image Processing (ICIP)*, pages 1369–1372, Hong Kong, Sep 2010b. (Cité page 82.)

L. Guo, N. Chehata, C. Mallet, and S. Boukir. Relevance of airborne lidar and image data for urban scene classification using random forests. *{ISPRS} International Journal of Photogrammetry and Remote Sensing*, 66(1) :56–66, Janvier 2011. (Cité pages 42, 44, 54, 56, 60, 63 et 78.)

I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using Support Vector Machines. *Machine Learning*, 46 :289–422, 2002. (Cité page 61.)

M. A. Hall and G. Holmes. Benchmarking attribute selection techniques for discrete class data mining. *IEEE Transactions on Knowledge and Data Engineering*, 15(6) :1437–1447, 2003. (Cité page 60.)

R. M. Haralick, K. Shanmugan, and I. Dinstein. Texture features for image classification. *IEEE T. on Systems, Man and Cybernetics*, pages 610–621, 1973. (Cité pages 57 et 62.)

H. Harzallah, F. Jurie, and C. Schmid. Combining efficient object localization and image classification. In *ICCV*, 2009. (Cité page 66.)

Martin Herold, Margaret E. Gardner, and Dar A. Roberts. Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9) :1907–1919, sep 2003. (Cité page 60.)

A. Hervieu, A. Le Bris, and C. Mallet. FUSION OF HYPERSPECTRAL AND VHR MULTISPECTRAL IMAGE CLASSIFICATIONS IN URBAN AREAS. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-3 :457–464, 2016. (Cité page 99.)

G.F. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1) :55–63, jan 1968. (Cité page 55.)

Jordi Inglada and Julien Michel. Qualitative Spatial Reasoning for High-Resolution Remote Sensing Image Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2) :599–612, Feb 2009. (Cité page 66.)

D. Jaskerniak, P.N.J. Lane, A. Robinson, and A. Lucieer. Extracting lidar indices to characterise multilayered forest structure using mixture distribution functions. *Remote Sensing of Environment*, 115 :573–585, 2010. (Cité page 73.)

K. Johansen, L. Arroyo, S. Phinn, C. Witte, R. Kauth, and G. Thomas. Comparison of geo-object based and pixel-based change detection of riparian environments using high spatial resolution multi-spectral imagery. *Photogrammetric Engineering and Remote Sensing*, 76(2) :123–136, 2010. (Cité page 67.)

S. Kandasamy, F. Tavin, A. Minghelli-Roman, S. Mathieu, Liu Weidong, F. Baret, and P. Gouton. Optimization of image parameters using a hyperspectral library application to soil identification and moisture estimation. In *Geoscience and Remote Sensing Symposium,2009 IEEE International, IGARSS 2009*, volume 3, pages 141–144, jul 2009. (Cité pages 56 et 60.)

K. Karantzalos and N. Paragios. Recognition-driven two-dimensional competing priors toward automatic and accurate building detection. *TGRS*, 47(1) : 133–144, 2009. (Cité page 66.)

M. Khodadadzadeh, J. Li, S. Prasad, and A. Plaza. Fusion of hyperspectral and lidar remote sensing data using multiple feature learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6) :2971–2983, June 2015. ISSN 1939-1404. (Cité page 54.)

S. Kluckner, T. Mauthner, P. Roth, and H. Bischof. Semantic classification in aerial imagery by integrating appearance and height information. In *ACCV*, pages 477–488, 2009. (Cité page 66.)

V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions via graph cuts. In *IEEE International Conference on Computer Vision*, pages 508–515, 2001. (Cité page 70.)

I. Kononenko, E. Simec, and M. Robnik-Sikonja. Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence*, 7(1) :39–55, 1997. (Cité page 60.)

M. Pawan Kumar, PHS. Torr, and A. Zisserman. OBJCUT : Efficient segmentation using top-down and bottom-up cues. *PAMI*, 14(1) :234–778, 2010. (Cité page 66.)

Camille Kurtz, Nicolas Passat, Anne Puissant, and Pierre Gançarski. *Hierarchical Segmentation of Multiresolution Remote Sensing Images*, pages 343–354. Ed. Soille, Pierre and Pesaresi, Martino and Ouzounis, Georgios K. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. ISBN 978-3-642-21569-8. (Cité page 69.)

L. Ladicky, P. Sturgess, K. Alahari, C. Russell, and PHS. Torr. What, Where & How Many ? Combining object detectors and CRFs. In *ECCV*, 2010. (Cité page 66.)

D. Larlus and F. Jurie. Combining appearance models and Markov Random Fields for category level object segmentation. In *CVPR*, 2008. (Cité page 66.)

P. Lassalle, J. Inglada, J. Michel, M. Grizonnet, and J. Malik. A scalable tile-based framework for region-merging segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 53(10) :5473–5485, Oct 2015. ISSN 0196-2892. (Cité page 71.)

A. Le-Bris, N. Chehata, X. Briottet, and N. Paparoditis. Contribution of texture and red-edge band for vegetated areas detection and identification. In *IEEE International Geoscience and Remote Sensing Symposium IGARSS 2013*, pages 4102–4105, 2013. (Cité pages 56 et 65.)

A. Le-Bris, N. Chehata, X. Briottet, and N. Paparoditis. Use intermediate results of wrapper band selection methods : A first step toward the optimization of spectral configuration for land cover classifications. In *IEEE WHISPERS*, Lausanne, Suisse, Jun 2014a. (Cité page 64.)

A. Le-Bris, N. Chehata, X. Briottet, and N. Paparoditis. Identify important spectrum bands for classification using importances of wrapper selection applied to hyperspectral data. In *IEEE IWCIM (International Workshop on Computational Intelligence for Multimedia Understanding)*, Paris, 2014b. (Cité page 64.)

A. Le-Bris, N. Chehata, X. Briottet, and N. Paparoditis. Extraction of optimal spectral bands using hierarchical band merging out of hyperspectral data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Volume XL-3/W3, pages 459–465, Oct 2015. (Cité page 64.)

A. Le-Bris, N. Chehata, X. Briottet, and N. Paparoditis. Hierarchically exploring the width of spectral bands for urban material classification. In *JURSE 2017 Joint Urban Remote Sensing Event*, March 2017. To appear. (Cité page 64.)

Arnaud Le-Bris. *Optimisation de la configuration dâun instrument superspectral aéroporté pour la classification : application au milieu urbain*. PhD thesis, ED MSTIC Université Paris-Est, 2015. (Cité pages 35, 39, 41, 56, 58, 60, 62 et 64.)

Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Y. Hardeberg. A constrained band selection method based on information measures for spectral image color visualization. *IEEE Transactions on Geoscience and Remote Sensing*, 49(12) :5104–5115, dec 2011. (Cité page 60.)

Christian Leistner, Amir Saffari, Jakob Santner, and Horst Bischof. Semi-Supervised Random Forests. In *IEEE International Conference on Computer Vision (ICCV)*, 2009. (Cité page 82.)

A. Levin and Y. Weiss. Learning to combine bottom-up and top-down segmentation. In *ECCV*, pages 581–594, 2006. (Cité page 66.)

Shijin Li, Hao Wu, Dingsheng Wan, and Jiali Zhu. An effective feature selection method for hyperspectral image classification based on genetic algorithm and Support Vector Machine. *Knowledge-based Systems*, 24 :40–48, 2011. (Cité pages 60 et 62.)

S. Liberge, B. Soheilian, N. Chehata, and N. Paparoditis. Extraction of vertical objects in 3D point clouds. In *ISPRS Technical Commission III Symposium on Photogrammetry Computer Vision and Image Analysis*, Paris, 1-3 Septembre 2010. (Cité pages 43 et 77.)

J. Liu and Y.H. Yang. Multiresolution color image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(7) :689–700, 1994. (Cité page 68.)

Y. Liu, L. Bian, Y. Meng, H. Wang, S. Zhang, Y. Yang, X. Shao, and B. Wang. Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68(0) :144–156, 2012. ISSN 0924-2716. (Cité pages 68 et 69.)

L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J. Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya. Hyperspectral pansharpening : A review. *IEEE Geoscience and Remote Sensing Magazine*, 3 (3) :27–46, Sept 2015. ISSN 2473-2397. (Cité page 54.)

J. Amorós López, E. Izquierdo Verdiguier, L. Gómez Chova, J. Mu noz MarÃ, J.Z. RodrÃguez Barreiro, G. Camps Valls, and J. Calpe Maravilla. Land cover classification of vhr airborne images for citrus grove identification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(1) :115–123, 2011. ISSN 0924-2716. (Cité page 67.)

M. Maire, S. Yu, and P. Perona. Object detection and segmentation from joint embedding of parts and pixels. In *ICCV*, 2011. (Cité page 66.)

C. Mallet, N. Chehata, and J-S. Bailly. *Méthodes de traitement de données LIDAR,*. Télédétection pour l'Observation des Surfaces Continentales, Tome 1 : Techniques de télédétection : principes et méthodes. ISTE Editeur, Elsevier, Système Terre â Environnement, 2016. (Cité pages ix, 42 et 57.)

G. Mallinis, N. Koutsias, M. Tsakiri-Strati, and M. Karteris. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(2) :237–250, 2008. ISSN 0924-2716. (Cité page 67.)

Adolfo Martínez-Usó, Filiberto Pla, José Martínez Sotoca, and Pedro García-Sevilla. Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12) :4158–4171, dec 2007. (Cité page 60.)

Ujjwal Maulik and Debasis Chakraborty. Learning with transductive {SVM} for semisupervised pixel classification of remote sensing imagery. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 77 :66–78, 2013. ISSN 0924-2716. (Cité page 82.)

F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8) :1778–1790, aug 2004. (Cité page 55.)

T. Melzer. Non-parametric segmentation of als point clouds using mean shift. *Journal of Applied Geodesy*, 1(3) :159–170, 2007. (Cité page 73.)

J. Minet, J. Taboury, M. Pealat, N. Roux, J. Lonnoy, and Y. Ferrec. Adaptive band selection snapshot multispectral imaging in the VIS/NIR domain. *Proceedings of SPIE the International Society for Optical Engineering*, 7835 :10, 2010. (Cité page 62.)

M. Dalla Mura, J.A. Benediktsson, B. Waske, and L. Bruzzone. Morphological attribute profiles for the analysis of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10) :pp 3747–3762, 2010. (Cité page 57.)

M. Neubert, H. Herold, and G. Meinel. Assessing image segmentation quality â concepts, methods and application. In Thomas Blaschke, Stefan Lang, and Geoffrey J. Hay, editors, *Object-Based Image Analysis*, Lecture Notes in Geoinformation and Cartography, pages 769–784. Springer Berlin Heidelberg, 2008. (Cité pages 68 et 69.)

Ali Ozgun Ok. Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts. *International Journal of Photogrammetry and Remote Sensing*, 86 :21–40, 2013. (Cité page 67.)

A. Ortiz and G. Oliver. On the use of the overlapping area matrix for image segmentation evaluation : A survey and new performance measures. *Pattern Recognition Letters*, 27(16) :1916–1926, 2006. ISSN 0167-8655. (Cité page 69.)

M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from digital elevation models using marked point processes. *International Journal of Computer Vision*, 72(2) :107–132, avril 2007. (Cité page 67.)

Julien Osman, Jordi Inglada, and Jean-FranÃ§ois Dejoux. Assessment of a markov logic model of crop rotations for early crop mapping. *Computers and Electronics in Agriculture*, 113 :234–243, 2015. ISSN 0168-1699. (Cité page 95.)

M. Pal and G.M. Foody. Feature selection for classification of hyperspectral data by svm. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5) :2297–2307, 2010. (Cité pages 55 et 78.)

S. Paris and F. Durand. A topological approach to hierarchical segmentation using mean shift. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Jun 2007. (Cité pages 72 et 76.)

Y. Pastol. Use of airborne lidar bathymetry for coastal hydrographic surveying : The French experience. *Journal of Coastal Research*, 62 :62–74, 2011. (Cité page 87.)

Moacir Pedroso, James Taylor, Bruno Tisseyre, Brigitte Charnomordic, and Serge Guillaume. A segmentation algorithm for the delineation of agricultural management zones. *Computers and Electronics in Agriculture*, 70(1) : 199–208, 2010. ISSN 0168-1699. (Cité page 67.)

Charlotte Pelletier, Silvia Valero, Jordi Inglada, Nicolas Champion, and GÃÂ©rard Dedieu. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, 187 :156–168, 2016. ISSN 0034-4257. (Cité page 78.)

Francesco Pirotti. Assessing a template matching approach for tree height and position extraction from lidar-derived canopy height models of pinus pinaster stands. *Forests*, 1(4) :194–208, 2010. (Cité page 73.)

V. Poulain, J. Inglada, M. Spigai, J. Y. Tourneret, and P. Marthon. High-resolution optical and sar image fusion for building database updating. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8) :2900–2910, Aug 2011. (Cité page 54.)

M. Segui Prieto and A.R. Allen. A similarity metric for edge images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10) :1265–1272, 2003. (Cité page 68.)

P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15 :1119–1125, nov 1994. (Cité pages 61 et 62.)

S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondance problem. In *IEEE International Conference on Computer Vision*, pages 492–502, 1998. (Cité page 70.)

R.E. Schapire, Y. Freund, P. Bartlett, and W.S. Lee. Boosting the margin : a new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5) :1651–1686, 1998. (Cité page 79.)

C. Senaras, M. Ozay, and F.T. Yarman Vural. Building detection with decision fusion. *JSTARS*, 6(3) :1295–1304, 2013. (Cité page 66.)

S.B. Serpico and G. Moser. Extraction of spectral channels from hyperspectral images for classification purposes. *IEEE Transactions on Geoscience and Remote Sensing*, 45(2) :484–495, feb 2007. (Cité page 60.)

J. Smeeckaert, C. Mallet, N. David, N. Chehata, and A. Ferraz. Large-scale classification of water areas using airborne topographic lidar data. *Remote Sensing of Environment*, 138 :134–148, 2013. ISSN 0034-4257. (Cité pages 42, 58, 70 et 88.)

D. Snow, P. Viola, and R. Zabih. Exact voxel occupancy with graph cuts. In *Computer Vision and Pattern Recognition Conf*, volume 1, pages 345–353, 2000. (Cité page 70.)

G. Sofia, J.-S. Bailly, N. Chehata, P. Tarolli, and F. Levavasseur. Comparison of pleiades and lidar digital elevation models for terraces detection in farmlands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(4) :1567–1576, 2016. ISSN 1939-1404. (Cité pages 44 et 77.)

J.M. Sotoca and P. Filiberto. Supervised feature selection by clustering using conditional mutual information-based distances. *Pattern Recognition*, 43(6) : 2068–2081, jun 2010. (Cité page 60.)

Victor F. Strîmbu and Bogdan M. Strîmbu. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 104 :30–43, 2015. (Cité page 73.)

A. Stumpf, N. Lachiche, J. P. Malet, N. Kerle, and A. Puissant. Active learning in the spatial domain for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5) :2492–2507, May 2014. (Cité pages 82 et 94.)

B. Thirion, B. Bascle, V. Ramesh, and N. Navab. Fusion of color, shading and boundary information for factory pipe segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 349–356, 2000. (Cité page 70.)

R. Tishirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, B-58(1) :267–288, 1996. (Cité page 61.)

P. Tittmann, S. Shafii, B. Hartsough, and B. Hamann. Tree detection, delineation, and measurement from lidar point clouds using ransac. In *SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems*, pages 1–13, Hobart, Australia, Oct 2011. Conference Secretariat. (Cité page 73.)

H. Torabzadeh, F. Morsdorf, and M.E. Schaepman. Fusion of imaging spectroscopy and airborne laser scanning data for characterization of forest ecosystems â a review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97 : 25–35, 2014. (Cité page 44.)

R. Trias-Sanz, G. Stamon, and J. Louchet. Using colour, texture, and hierarchial segmentation for high-resolution remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(2) :156 – 168, 2008. (Cité pages 67 et 68.)

D. Tuia, F. Ratle, F. Pacifici, M.F. Kanevski, and W.J. Emery. Active learning methods for remote sensing image classification. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(7) :2218–2232, July 2009. (Cité page 82.)

D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Munoz-Mari. A survey of active learning algorithms for supervised remote sensing image classification. *Selected Topics in Signal Processing, IEEE Journal of*, 5(3) :606–617, June 2011. (Cité page 82.)

R. Unnikrishnan, J.F. Lalonde, N. Vandapel, and M. Hebert. Scale selection for geometric fitting in noisy point clouds. *International Journal on Computational Geometry and Applications*, 20(5) :pp 543–575, 2010. (Cité page 58.)

F. VanCoillie, T. Novack, L. M. G. Fonseca, and H. J. R. Kux. Quantitative segmentation evaluation for large scale mapping purposes. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-4/C1, 2008. (Cité pages 68 et 69.)

F.M.B. VanCoillie, L.P.C. Verbeke, and R.R. De Wulf. Feature selection by genetic algorithms in object-based classification of ikonos imagery for forest mapping in flanders, belgium. *Remote Sensing of Environment*, 110(4) :476–487, 2007. (Cité page 68.)

M. C. Vanegas, I. Bloch, and J. Inglada. Alignment and parallelism for the description of high-resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6) :3542–3557, June 2013. (Cité page 66.)

O. Veksler. Image Segmentation by Nested Cuts. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 339–344, 2000. (Cité page 70.)

G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald. A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5) : 2565–2586, May 2015. (Cité page 54.)

Y. Wang, Y.S. Soh, and H. Shultz. Individual tree crown segmentation in aerial forestry images by mean shift clustering and graph-based cluster merging. *International Journal of Computer Science and Network Security*, 6(11) :40–45, 2006. (Cité page 71.)

Bernd-Michael Wolf and Christian Heipke. Automatic extraction and delineation of single trees from remote sensing data. *Machine Vision and Applications*, 18(5) :317â–330, 2007. ISSN 1432-1769. (Cité page 73.)

W. Xiao, B. Vallet, K. Schindler, and N. Paparoditis. Street-side vehicle detection, classification and change detection using mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114 :166–178, apr 2016. (Cité page 73.)

He Yang, Qian Du, and Genshe Chen. Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2) :544–554, apr 2012. (Cité page 62.)

Hua-chao Yang, Shu-bi Zhang, Ka-zhong Deng, and Pei-jun Du. Research into a feature selection method for hyperspectral imagery using PSO and SVM. *Journal of China University of Mining and Technology*, 17(4) :473–478, dec 2007. (Cité page 62.)

M.M. Sidi Youssef, C. Mallet, N. Chehata, A. Le Bris, and A. Gressin. Combining top-down and bottom-up approaches for building detection in a single very high resolution satellite image. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2014*, pages 4820–4823, 2014a. (Cité pages 74 et 102.)

M.M. Sidi Youssef, C. Mallet, N. Chehata, A. Le Bris, and A. Gressin. Détection de bâtiments à partir dâune image satellitaire par combinaison dâapproches ascendante et descendante. In *RFIA'14 (19 ème congrès national sur la Reconnaissance de Formes et l'Intelligence Artificielle)*, Rouen, Jun 2014b. (Cité pages 67 et 74.)

Mohamed Mahmoud Sidi Youssef. Amélioration de la segmentation d'images aériennes par reconnaissance de formes. Master's thesis, Sup'Com (Ecole Supérieure des Communications de Tunis), 2013. (Cité pages 76 et 101.)

H. Zhang, J.E. Fritts, and S.A. Goldman. Image segmentation evaluation : A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2) :260–280, 2008. (Cité page 68.)

J. Zhang. Multi-source remote sensing data fusion : status and trends. *International Journal of Image and Data Fusion*, 1(1) :5–24, 2010. (Cité page 54.)

Liangpei Zhang, Yanfei Zhong, Bo Huang, Jianya Gong, and Pingxiang Li. Dimensionality reduction based on clonal selection for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12) :4172–4186, dec 2007. (Cité page 60.)

Min-Ling Zhang and Zhi-Hua Zhou. Exploiting unlabeled data to enhance ensemble diversity. *Data Min. Knowl. Discov.*, 26(1) :98–129, jan 2013. (Cité page 95.)

Q. Zhang, V.P. Pauca, and R.J. Plemmons et al. Detecting objects under shadows by fusion of hyperspectral and lidar data : a physical model approach. In *Proceedings of 5th Workshop Hyperspectral Image Signal Processing Evolution in Remote Sensing*, pages 1–4, 2013. (Cité page 44.)

Y.J. Zhang. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8) :1335–1346, 1996. (Cité page 68.)

S. Zhou, J.-P. Zhang, and B.-K. Su. Feature selection and classification based on ant colony algorithm for hyperspectral remote sensing images. In *Proc. of the 2nd International Congress on Image and Signal Processing (CISP'09)*, pages 1–4, oct 2009. (Cité page 62.)

J. Zhu, S. Rosset, T. Hastie, and R. Tibshirani. 1-norm support vector machines. *Advances in neural information processing systems*, 16(1) :49–56, 2004. (Cité page 61.)

Xiaojin Zhu. Semi-supervised learning literature survey. Technical Report 1530, Computer Sciences, University of Wisconsin-Madison, 2005. (Cité page 82.)

L. Zhuo, J. Zheng, F. Wang, X. Li, A. Bin, and J. Qian. A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7) :397–402, jul 2008. (Cité pages 60 et 62.)