

Analyse et traitement d'ondes lidar pour la cartographie et la reconnaissance de formes : application au milieu urbain

Lidar waveform analysis and processing for cartography and pattern recognition : application to urban areas

Clément Mallet¹Adrien Chauve^{1,2}Frédéric Bretar¹

¹ Laboratoire MATIS - Institut Géographique National,
2-4 av. Pasteur, 94165 Saint-Mandé

² UMR TETIS Cemagref/Cirad/ENGREF, Maison de la Télédétection
500, rue J.F. Breton 34093 Montpellier Cedex 5
{clement.mallet, adrien.chauve, frederic.bretar}@ign.fr

Résumé

Toute onde lidar rétrodiffusée par la surface terrestre contient des informations sur les cibles atteintes ayant contribué à la forme de l'onde. Les systèmes lidar capables de numériser l'intégralité des signaux retour sont apparus récemment et permettent le traitement a posteriori de ces profils altimétriques. Nous présentons dans cet article une méthode d'analyse puis de traitement des ondes lidar dans un contexte de cartographie automatique. Tout d'abord, nous montrons que l'analyse fine des ondes permet une densification des nuages de points 3D. Dans un second temps, le traitement a posteriori des signaux conduit à leur modélisation sous forme paramétrique. Nous proposons alors une méthode de reconnaissance de formes appliquée au milieu urbain. Une classification supervisée par Séparateurs à Vaste Marge est ainsi employée pour prendre en compte les caractéristiques des échos extraits lors de la phase de traitement. Les résultats montrent que la segmentation d'une zone urbaine en classes bâti, végétation, sol naturel et sol artificiel est possible à partir des ondes lidar seulement.

Mots Clef

Onde lidar, estimation paramétrique, modélisation, classification supervisée, milieu urbain.

Abstract

Each lidar waveform backscattered by the Earth's surface provide information about the reflecting targets giving its shape to the signal. Lidar systems are recently able to digitize the full-waveform and enable to post process these altimetric profiles. In this article, we present how lidar waveforms can be analysed and processed for automatic cartographic purposes. First, fine waveform analysis allows to densify 3D point clouds. Besides, parametric signal mo-

delling is carried out. Then, a image-based pattern recognition method for urban areas is proposed. A supervised classification using Support Vector Machines is therefore performed to use jointly the parameters extracted from the post processing step. Results show that it is possible to segment urban areas in building, vegetation, natural ground and artificial ground labels using only lidar waveforms.

Keywords

Lidar waveform, parametric estimates, modelling, supervised classification, urban areas.

1 Introduction

Depuis quelques années, la technique de télémétrie laser, ou lidar, adaptée au contexte aérien fournit une mesure de distance entre le vecteur aérien et la surface terrestre. Le calcul de la trajectographie du vecteur aérien fondée sur le couplage d'une centrale inertielle (INS) et d'un GPS différentiel permet de transformer ces mesures de distance en nuages de points 3D [1]. La précision altimétrique de la mesure est particulièrement élevée (< 0,15 m) en lidar topographique et permet également d'acquérir plusieurs points de mesure le long de l'onde réfléchie.

Un système lidar aéroporté émet des impulsions laser à haute fréquence. La forme de l'onde rétrodiffusée collectée par le système dépend de la répartition spatiale et des caractéristiques des objets atteints par le faisceau laser. Il existe deux types de systèmes lidar :

- les systèmes traditionnels "multi-échos" détectent seulement la position de maxima d'intensité pour générer de manière automatique des points 3D ;
- les systèmes dit à "retour d'onde complète", plus récents, enregistrent l'intégralité de l'onde.

Cet article explore les potentialités de l'analyse des données issues des systèmes à "retour d'onde complète" pour la cartographie et la reconnaissance de formes.

Le premier système opérationnel, LVIS (Laser Vegetation Imaging Sensor), satellitaire, est apparu en 1999 [2]. Conçus pour décrire la structure verticale de la canopée, LVIS et ses successeurs ont démontré la pertinence de l'enregistrement de l'intégralité des ondes lidar pour caractériser les surfaces forestières, mesurer la topographie terrestre y compris sous la canopée et cartographier grossièrement les zones d'occupation des sols [3, 4]. Le premier système aéroporté commercial est opérationnel depuis 2004. L'exploitation des ondes lidar pour la cartographie à moyenne et petite échelles est donc possible depuis peu.

De nombreuses études ont été menées pour analyser et traiter les ondes enregistrées par ces systèmes. La première étape est le traitement de l'onde retour pour extraire plus qu'une simple mesure de distance (*ie* plus qu'un point 3D). Pour cela, les signaux sont ajustés par une somme de fonctions gaussiennes, les paramètres étant estimés par maximum de vraisemblance [5] ou par moindres carrés non linéaires (MCNL) [6]. Il a été montré que cette approximation était la plus simple et la plus pratique pour traiter les ondes lidar [7]. Ensuite, des recherches sont menées sur la classification de scènes urbaines (discrimination de matériaux, détection de pans de toits, de la végétation) [8, 9] et l'estimation de paramètres forestiers à grande échelle [3]. Elles ont montré la nécessité d'une connaissance théorique de l'influence des cibles atteintes sur la forme de l'onde lidar retour pour espérer approfondir ces recherches.

Le but de l'étude présentée dans cet article est double. D'une part, nous montrons que l'analyse fine des ondes lidar permet une densification des nuages de points 3D. Plus d'informations altimétriques sont extraites, utiles pour la caractérisation des paysages survolés. Ensuite, le traitement a posteriori permet de modéliser l'onde retour sous forme paramétrique. Nous proposons alors une méthode de discrimination des paysages appliquée au milieu urbain dans un contexte de cartographie automatique en s'appuyant sur un nuage de points 3D à la fois densifié et paramétrisé.

La première partie de cet article décrit les systèmes lidar à retour d'onde complète ainsi que les données générées. Puis nous traitons le problème de l'ajustement paramétrique des ondes retour en proposant plusieurs fonctions mathématiques convenables (partie 3). Ensuite, une méthodologie de reconnaissance d'objets à partir des ondes traitées est décrite (partie 4). Une image fondée sur les premiers échos détectés dans les ondes lidar est segmentée finement. Un vecteur de paramètres discriminants est alors calculé pour chaque région. La méthode de classification supervisée par Séparateurs à Vaste Marge (SVM) est alors utilisée pour discriminer quatre grandes classes : bâtiments, végétation haute, sol naturel (herbe, terre battue...) et sol artificiel (rues, parkings). Enfin, les résultats sur un paysage urbain sont présentés dans la partie 5. Nous

concluons cet article en mentionnant différentes perspectives de recherche sur ces travaux.

2 Présentation des données lidar

2.1 Intérêt de l'enregistrement de signaux lidar

Le principe de fonctionnement d'un système lidar aéroporté consiste en l'émission d'impulsions laser de formes gaussiennes (d'une durée de 5 à 10 ns) et de fréquence temporelle élevée (~ 100 kHz). Chaque impulsion laser interagit avec les objets présents le long du faisceau à différentes couches altimétriques. Le signal enregistré par le détecteur est donc composé de la somme des rétrodiffusions de chacun des objets (*cf.* figure 1).

Les systèmes lidar dit "multi-échocs" détectent en temps-

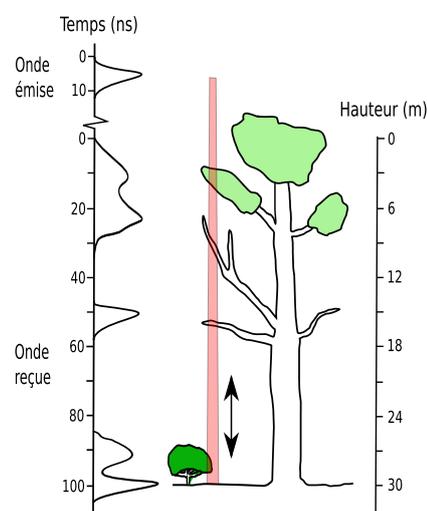


FIG. 1 – Ondes laser émise et reçue sur une zone de végétation complexe. Tous les objets suffisamment résolus spatialement contribuent à la forme de l'onde.

réel un certain nombre de maxima d'intensité (pouvant aller jusqu'à six) dans l'onde rétrodiffusée suivant une méthode prédéfinie, le plus souvent inconnue des utilisateurs [10]. Ces maxima sont appelés "échocs". Le produit final de tels systèmes est alors un nuage de points 3D géoréférencé décrivant la topographie terrestre.

Cependant, les répartitions altimétriques et spectrales des objets sont rarement homogènes à l'intérieur du cône de diffraction laser. L'impulsion émise est donc déformée lors de sa réflexion sur la surface terrestre. La forme parfois complexe de l'onde contient des informations sur la radiométrie et la géométrie des cibles atteintes. Les facteurs d'influence sont nombreux (angle d'émission, couleur, type de matériau, rugosité, forme, pente) et de nombreuses études sont menées pour connaître leurs influences spécifiques [11]. En ne gardant finalement qu'un nuage de

Système	LMS-Q560
Hauteur de vol	500 m
Longueur de l'impulsion	5 ns
Fréquence d'acquisition	100 kHz
Fréquence de numérisation	1 ns
Résolution du système	0,6 m
Densité d'impulsions	2,5/m ²

TAB. 1 – Caractéristiques techniques du système RIEGL© LMS-Q560 et paramètres d'acquisition des données.

points 3D, ces informations sont perdues.

Les systèmes lidar à retour d'onde complète en plus de conserver toute l'information disponible, autorisent le traitement des données a posteriori (par rapport au traitement temps-réel des systèmes multi-échos [12]). Cela permet donc de mieux contrôler l'extraction des échos et d'améliorer la connaissance de la distribution altimétrique des cibles sur le trajet de l'onde.

2.2 Ondes lidar traitées

Les données utilisées dans cette étude sont enregistrées par un système lidar à retour d'onde complète. Une onde lidar est en fait un profil altimétrique composé d'une ou de plusieurs séquences de 18 m, échantillonnées tous les 0,3 m. Le nombre de séquences dépend de la hauteur totale des objets présents sur le trajet du faisceau laser. En pratique, les signaux lidar traités sont donc des séquences de 60 échantillons (voir figure 2).

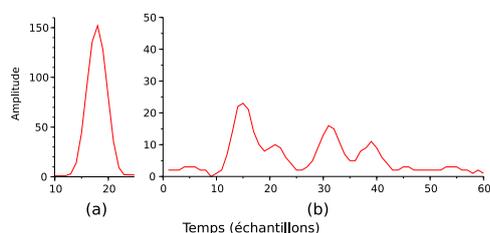


FIG. 2 – (a) Onde émise par le lidar. (b) Onde rétrodiffusée par une zone de végétation. Le bruit de fond y est relativement faible.

Les données de cette étude ont été acquises en septembre 2006 par la société Toposys© avec un système lidar RIEGL© sur la ville de Biberach en Allemagne. La zone acquise comprend à la fois des zones résidentielles, industrielles ainsi que des zones de forêt dense. Le tableau 1 présente les principales caractéristiques du système utilisé.

Chaque onde lidar est composée d'une ou de deux séquences de 60 échantillons, c'est-à-dire un profil altimétrique de 18 m, ou 36 m. Pour chaque impulsion laser, le signal émis (60 échantillons également) est fourni ainsi que les échos détectés par le système temps-réel embarqué lors

de l'acquisition des données (figure 2).

3 Modélisation des ondes rétrodiffusées

Le traitement des ondes lidar consiste en premier lieu en leur décomposition en une somme d'échos distincts afin de caractériser les différentes cibles rencontrées sur le trajet du faisceau laser. Ce processus de modélisation des ondes lidar est crucial pour densifier le nuage de points [3] et pour étudier la morphologie des réponses spécifiques des objets dans un objectif de classification [13].

D'un point de vue méthodologique, des stratégies d'estimation paramétrique ou non-paramétrique conviennent. L'approche paramétrique consiste à choisir un modèle mathématique donné où chaque classe est représentée par une fonction analytique et formule le problème d'estimation de la fonction comme un problème d'estimation de paramètres. Plusieurs méthodes permettent cette estimation, qu'il s'agisse d'une méthode par maximum de vraisemblance, moindres carrés non linéaires ou en utilisant la "méthode des moments". A contrario, l'estimation non paramétrique n'est fondée sur aucun modèle analytique spécifique mais fournit une plus grande flexibilité à l'ajustement aux données (réseau de neurones, splines ou fenêtres de Parzen) [14].

3.1 Méthodologie

L'approche choisie dans cet article est paramétrique : la contribution des cibles est modélisée sous forme analytique par une fonction modèle. L'objectif est d'extraire pour chaque onde autant d'échos que possible ainsi que des informations sur leur morphologie. Cela se fait en deux grandes étapes : tout d'abord, le nombre d'échos est estimé ainsi que les valeurs initiales des paramètres de la fonction. Ensuite, ces paramètres sont estimés par optimisation. Il a été montré que pour les ondes lidar les algorithmes d'*Expectation-Maximization* (estimation par maximum de vraisemblance) [5] ou de Levenberg-Marquardt [6, 7] donnent de bons résultats. Le choix de la méthode d'optimisation n'est donc pas crucial. Néanmoins, la qualité du résultat dépend fortement de l'estimation des paramètres initiaux. Par conséquent, dans cette partie nous mettons l'accent sur l'estimation initiale de la position des maxima du signal puis proposons une modélisation des ondes lidar avec différentes fonctions pour améliorer l'ajustement du signal.

Une onde lidar est composée d'une séquence de points uniformément espacés $\{(x_i, y_i)\}_{i=1, \dots, N}$. On cherche à décomposer chaque séquence en une somme de composantes représentant les cibles situées dans le trajet de l'onde sous la forme

$$y = f(x) = \sum_{j=1}^n f_j(x) + b \quad (1)$$

où n est le nombre de composantes, f_j une fonction donnée

qui peut être gaussienne, log-normale ou gaussienne généralisée (voir partie 3.3) et b le bruit de fond des mesures (négligé dans cette étude).

Pour chaque séquence, le signal est tout d'abord seuillé pour éliminer le bruit et éviter ainsi la détection de maxima ne correspondant à des cibles atteintes. Ensuite, une détection simple de maxima locaux est utilisée pour estimer le nombre et la position des composantes. Les autres paramètres du modèle choisi sont fixés constants. Après un premier ajustement par MCNL, une détection des pics manquants est alors menée (échos se recouvrant, voir figure 3).

3.2 Détection des échos

La détection simple des maxima d'intensité dans l'onde se fait par la recherche des passages par zéro de la dérivée première du signal seuillé. L'algorithme prend en compte le nombre minimal d'échantillons séparant deux échos (résolution spatiale du système). Une méthode par MCNL avec l'algorithme de Levenberg-Marquardt est choisie pour calculer l'ajustement. La qualité des résultats est évaluée par le critère

$$\xi = \frac{1}{N - p} \sum_{k=1}^N (y_k - f(x_k))^2 \quad (2)$$

où le numérateur est la somme des résidus entre l'onde observée et le signal modélisé, N le nombre d'échantillons et p le nombre de paramètres du modèle.

Dans le cas d'échos se recouvrant, la détection simple n'est pas suffisante car chaque écho ne correspond pas forcément à un maximum local du signal. Une seconde détection sur la différence entre l'onde initiale et l'ajustement calculé est effectuée. Si un pic est trouvé, un nouvel ajustement est mené. Le paramètre de qualité ξ est comparé au précédent et cette étape est répétée jusqu'à ce que le facteur ξ arrête de décroître.

3.3 Choix des fonctions analytiques

Chaque onde laser émise est considérée comme étant de forme gaussienne avec une largeur calibrée. L'onde rétro-diffusée enregistrée par le lidar est alors la convolution de cette distribution gaussienne par une fonction "terrain" qui dépend des objets atteints par le faisceau laser. Il a été montré que si la distribution altimétrique des objets dans le cône de diffraction suit un modèle gaussien, alors l'onde lidar peut être approchée par une somme de fonctions gaussiennes [4]. Wagner [7] a par ailleurs montré que plus de 98 % des ondes lidar issues du système RIEGL© pouvaient être estimées par une somme de gaussiennes.

Cependant, cette approximation n'est pas forcément satisfaisante. Selon le système lidar, l'onde transmise n'est pas toujours gaussienne et peut être déformée (asymétrique, aplatie ou pointue). De plus, la distribution altimétrique gaussienne des objets dans le cône de diffraction laser n'a été montrée que pour des systèmes lidar à large empreinte au sol (15-100 m). Pour des systèmes à petite empreinte (0,5 m dans notre cas), on ne peut faire cette hypothèse et la modélisation par une somme de fonctions gaussiennes

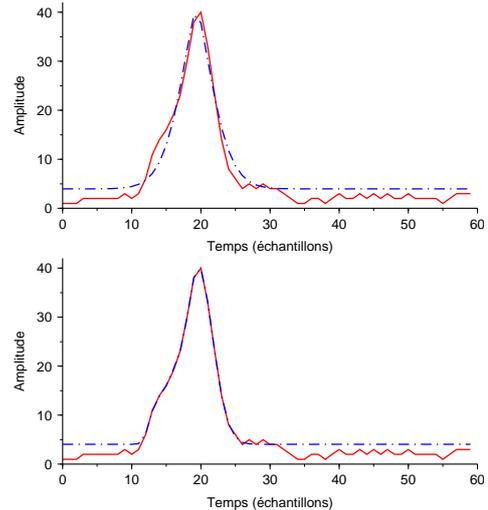


FIG. 3 – Exemple d'onde lidar complexe. Donnée lidar en trait rouge continu et résultat de l'ajustement en pointillés bleu. Les données ont été seuillées à 4 avant la détection des échos. **Haut** : Ajustement avec uniquement une détection d'échos grossière (un écho trouvé). **Bas** : Ajustement avec une détection fine (deux échos).

est inexacte.

Il est donc intéressant d'étendre le traitement classique des ondes lidar à des modèles paramétriques différents et plus complexes. Cela permet d'une part de mieux estimer les ondes lidar mais également d'extraire plus d'informations des signaux bruts. Les fonctions log-normale et gaussienne généralisée sont ainsi utilisées pour modéliser les signaux asymétriques (figure 4a) et déformés. Les modèles gaussien (G), log-normal (L) et gaussien généralisé (GG) ont les expressions analytiques suivantes (voir la figure 4 pour leurs tracés) :

$$f_{G,j}(x) = a_j \exp\left(-\frac{(x - \mu_j)^2}{2\sigma_j^2}\right) \quad (3)$$

$$f_{L,j}(x) = a_j \exp\left(-\frac{(\ln(x - s_j) - \mu_j)^2}{2\sigma_j^2}\right) \quad (4)$$

$$f_{GG,j}(x) = a_j \exp\left(-\frac{|x - \mu_j|^{\alpha_j}}{2\sigma_j^2}\right) \quad (5)$$

La paramètre a_j correspond à l'amplitude de l'écho et σ_j à sa largeur. La position de l'écho (distance au capteur) est donnée par μ_j pour les gaussiennes simple et généralisée, et par $s_j + \exp(\mu_j - \sigma_j^2)$ pour la log-normale.

L'observation des ondes lidar sur la zone d'acquisition montre que la plupart des ondes de forme asymétrique sont en fait la contribution d'objets distincts trop proches spatialement pour qu'un troisième point d'inflexion puisse être détecté. Ces cas sont bien modélisés par la fonction log-normale.

Le modèle gaussien généralisé permet de simuler à la fois

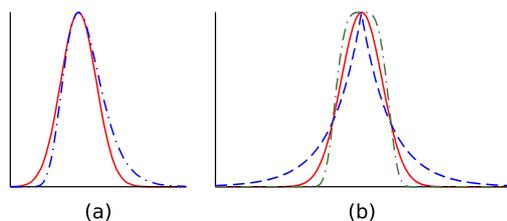


FIG. 4 – (a) Comparaison entre les modèles gaussien (ligne continue) et log-normal (pointillés). (b) La fonction gaussienne généralisée : $\alpha=1$ (tirets), $\alpha=\sqrt{2}$ (ligne continue) et $\alpha=2$ (pointillés).

des formes gaussiennes ($\alpha = \sqrt{2}$), des formes pointues ($1 \leq \alpha < \sqrt{2}$, $\alpha = 1$ donne la fonction de Laplace) et des formes aplaties au sommet ($\alpha > \sqrt{2}$, cf. figure 4b). Par conséquent, l'utilisation d'une telle fonction devrait améliorer la qualité de l'ajustement dans tous les cas. En pratique, le nombre d'estimations qui divergent augmentent aussi. Cela est dû au plus grand nombre de degrés de libertés de la fonction (plus de paramètres à estimer) et à l'expression plus complexe du gradient associé [15]. Le paramètre α est potentiellement intéressant pour l'analyse d'ondes lidar et la classification de nuages de points puisqu'il fournit une information supplémentaire sur la forme de l'écho (cf. partie 5.2).

4 Classification

4.1 Approche proposée

Pour étudier les potentialités des informations extraites des ondes lidar dans un problème de reconnaissance de formes appliqué au milieu urbain, nous avons adopté la stratégie suivante :

- création d'images de données lidar dans la géométrie d'acquisition ;
- segmentation ;
- classification supervisée par Séparateurs à Vaste Marge (SVM).

Quatre classes ont été choisies pour décrire le milieu urbain : bâtiment, végétation haute, sol naturel et sol artificiel.

Chaque pixel des images de données lidar décrit la région de l'espace illuminée par un unique faisceau laser. Autant d'images sont créées qu'il y a de paramètres extraits des ondes. La figure 5 montre un exemple d'une telle image.

La segmentation est fondée sur le formalisme de segmentation multi-échelles décrit dans [16]. À partir d'une segmentation par Ligne de Partage des Eaux, une hiérarchie est construite en fusionnant de manière itérative les deux régions les plus similaires d'un point de vue énergétique [17]. Cette méthode générale permet, en partant d'une segmentation quasi-pixellaire, de choisir l'échelle d'analyse souhaitée d'une image par une coupe à un niveau donné dans la hiérarchie. Pour notre application, l'utilisation de cette mé-

thode permet de sélectionner un niveau de partitionnement respectant les formes que l'on cherche à reconnaître (i.e. l'échelle d'un bâtiment ou d'un arbre). L'image segmentée est une image à trois canaux : amplitude, forme (paramètre α) et altitude du premier écho. Avec cette approche hiérarchique, l'image labellisée est finalement composée à la fois de petites régions, du niveau de la segmentation par Ligne de Partage des Eaux, et de grandes régions relativement planes, homogènes radiométriquement et du point de vue de la forme des échos.

Les segments résultants sont alors classés par Séparateurs à Vaste Marge (SVM). La méthode SVM est bien adaptée pour traiter des données de grande dimension (la complexité de l'algorithme ne dépend pas de la dimension des données) et très exploitée pour des problèmes de reconnaissance de forme en télédétection. L'utilisation d'un simple vecteur d'attributs permet la fusion de données issues de sources hétérogènes [18]. L'approche proposée dans cette article se fonde uniquement sur des données issues du traitement d'ondes lidar.

4.2 Séparateurs à Vaste Marge

Dans cette partie, nous présentons brièvement la méthode de classification par SVM. Soient \mathcal{D} l'espace des données et \mathcal{Y} l'espace des classes ($\mathcal{Y} = \{-1, 1\}$ par exemple). $x_i \in A$ représentent les données d'apprentissage (sous-ensemble de $\mathcal{D} \times \mathcal{Y}$) et ω le classifieur. Le but de la classification supervisée est de construire à partir des données d'apprentissage un classifieur robuste. Dans le cadre des SVM, il s'agit de trouver un hyperplan permettant la séparation de l'ensemble d'apprentissage de sorte à ce que tous les points d'une même classe soient du même côté de l'hyperplan [19]. Ainsi dans le cas d'un classifieur linéaire $\omega(x) = \mathbf{w} \cdot x - \theta$, où \mathbf{w} est le vecteur normal à l'hyperplan et θ le biais, nous imposons après normalisation :

$$\forall (x, y) \in A, \quad y \times (\mathbf{w} \cdot x - \theta) > 1 \quad (6)$$

Cela signifie la définition de deux hyperplans, parallèles au classifieur, cherchant à maximiser la distance entre l'hyperplan et les points les plus proches de chaque classe (distance définie comme $2/\|\mathbf{w}\|$). Il s'agit d'un problème d'optimisation cherchant à trouver le classifieur qui maximise la distance euclidienne entre les hyperplans (la marge) sous les contraintes définies dans l'équation 6. Dans la plupart des cas, il n'est pas possible de trouver un classifieur linéaire robuste à partir des données d'apprentissage. Une solution est alors de projeter les points d'apprentissage dans un espace \mathcal{H} où le problème est linéairement séparable. On introduit pour cela une fonction $\Phi : \mathcal{D} \rightarrow \mathcal{H}$ puis on applique la même méthode d'optimisation de la marge. La formulation du problème est alors la suivante :

$$\max_{\alpha} \sum_{i=1}^{\dim \mathcal{D}} \alpha_i - \sum_{i=1}^{\dim \mathcal{D}} \sum_{j=1}^{\dim \mathcal{D}} \alpha_i \alpha_j y_i y_j (\Phi(x_i) | \Phi(x_j)) \quad (7)$$

sous la contrainte : $\sum_{i=1}^{\dim \mathcal{D}} \alpha_i y_i = 0$ où les α_i sont les paramètres de Lagrange-Kuhn-Tucker et $(. | .)$ est le produit

scalaire euclidien dans \mathcal{H} .

Il s'agit désormais d'un problème d'optimisation convexe. La solution optimale α^* donne les paramètres optimaux de classification w^* et θ^* et finalement le classifieur optimal est donné par :

$$\omega^*(x) = \text{signe} \left(\sum_{i=1}^{\dim D} \alpha_i^* y_i (\Phi(x_i) | \Phi(x) - \theta^*) \right) \quad (8)$$

La recherche de la fonction Φ est très difficile. En pratique, la projection des points x_i dans \mathcal{H} se fait de manière implicite à travers la définition d'une fonction noyau $K : \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}$ avec $K(x_i, x_j) = (\Phi(x_i) | \Phi(x_j))$. La connaissance de K est suffisante pour calculer le classifieur optimal.

Le noyau choisi est le noyau gaussien dont la forme générique est $K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\gamma^2})$. Étant donné que la fonction de changement d'espace induit par ce noyau fait correspondre une fonction continue à chaque exemple, le nouvel espace de représentation \mathcal{H} est de dimension infinie. Une solution peut donc être trouvée [21]. Les deux paramètres de notre problème, γ , qui règle le degré de similarité de la classification sur les données d'apprentissage, et C le compromis entre maximisation de la marge et minimisation de l'erreur d'apprentissage sont calculés par validation croisée sur un échantillon des données test (différent des données d'apprentissage).

La classification finale est affinée en introduisant tout d'abord une classe de rejet. Les régions étiquetées non-classées sont celles dont aucune classe ne ressort très majoritaire par rapport aux autres. D'autre part, les régions de petite taille strictement incluses dans une autre de classe différente sont assimilées à ces dernières par analyse du contexte : une petite région étiquetée "bâtiment" incluse dans une région "sol naturel" sera par exemple intégrée à cette dernière.

Nous définissons enfin un critère de qualité, nommé précision de la classification, défini par :

$$\rho = \frac{\sum_{i=1}^{\dim \mathcal{Y}} A_{i,i}}{\sum_{i=1}^{\dim \mathcal{Y}} \sum_{j=1}^{\dim \mathcal{Y}} A_{i,j}} \in [0,1] \quad (9)$$

où $A_{i,j}$ correspond au nombre de segments classés dans j appartenant à la région i .

4.3 Sélection des attributs discriminants

Le vecteur d'attributs proposé pour chaque segment est composé de six éléments : la différence d'altitude moyenne entre le premier et le dernier écho de chaque onde $\mu_{\Delta z}$, la plus petite valeur propre λ_z de la matrice de variance-covariance d'altitude des premiers échos dans la région (résidus au plan estimé par moindres carrés à partir des points laser) et les valeurs moyennes de l'entropie, de l'intensité, de la largeur et de la forme du premier écho des ondes lidar incluses dans la région, respectivement $\mu_S, \mu_A, \mu_\sigma, \mu_\alpha$. Seul le premier écho est choisi pour calculer ces attributs

Attribut	$\mu_{\Delta z}$	λ_z	μ_S	μ_A	μ_σ	μ_α
Bât.	0	$\simeq 0$	var.	var.	moy.	var.
Vég. hte	forte	forte	forte	moy.	forte	var.
Sol nat.	0	$\simeq 0$	faible	faible	var.	$\geq \sqrt{2}$
Sol art.	0	$\simeq 0$	faible	var.	var.	$\simeq \sqrt{2}$

TAB. 2 – Valeurs empiriques des attributs discriminants du SVM sur les quatre classes testées (Bât. : bâtiment, Vég. hte : végétation haute, var. : variable (fort ou faible), moy. : moyenne).

car l'image sur-segmentée à labelliser est une image générée à partir des paramètres du premier écho de chaque onde.

L'entropie S est une mesure locale de structuration ici altimétrique. À partir de l'image d'altitude du premier écho, on mesure le degré de structuration au voisinage de chaque pixel. La mesure repose sur la répartition statistique des directions de gradient dans un voisinage centré sur un point. Les directions sont mesurées modulo $\pi/2$ et échantillonnées en 16 classes. Un histogramme des directions du gradient (de norme suffisamment grande) dans un disque de diamètre D est généré et on calcule la probabilité d'apparition p_i de chaque direction θ_i au sein du disque. Si le nombre de valeurs stockées est trop faible, le pixel est classé "homogène" ($S=0$). Sinon on calcule l'entropie de la densité de probabilité des directions $S = -\sum_{0 \leq i \leq 15} p_i \log p_i$.

$\mu_{\Delta z}$ et λ_z sont utilisés pour discriminer les zones de végétation dans l'image. μ_S différencie la végétation et certains bâtiments du sol. En effet, l'entropie est calculée sur une image d'altitude et non d'intensité. Les zones homogènes sont les zones à faibles variations altimétriques. Les toits de bâtiments surmontés par exemple de superstructures auront une entropie forte. Cependant, si l'étape de sur-segmentation a séparé toits et superstructures alors les zones de toits auront une entropie faible, c'est pourquoi d'autres attributs discriminants sont nécessaires. Nous avons pour cela introduit les trois paramètres issus de l'étape de modélisation.

Les valeurs théoriques des attributs pour chacune des quatre classes de l'étude sont décrites dans le tableau 2.

D'autres attributs ont été testés comme le nombre moyen d'échos par onde, la texture altimétrique par région ainsi que des paramètres de forme (compacité...) mais se sont redondants dans le processus de classification.

5 Résultats et discussions

5.1 Modélisation des signaux retour

Détection des échos.

La détection fine d'échos dans les ondes lidar par une recherche itérative du nombre de composantes permet de densifier le nuage de points 3D final jusqu'à plus de 50 % sur des zones de forêts par rapport au nuage produit de

Zone	Zone entière	Vég. dense	Vég.	Résidentiel	Centre-ville
Nb profils	2027547	70074	23368	93690	66264
Nb échos Supp.(%)	2903976 24	147218 55	46246 51	120813 9	85520 10
ΔH_p (m)	-0,13	-0,42	-0,34	-0,04	-0,04
ΔH_d (m)	0,36	1,58	1,36	0,07	0,05

TAB. 3 – Statistiques sur l'extraction d'échos dans les ondes lidar sur différentes zones test. Les chiffres sont donnés pour le modèle gaussien (Vég. : végétation, Supp. : points supplémentaires, ΔH_p : ΔH premier écho et ΔH_d : ΔH dernier écho).

manière automatique par le système RIEGL© (voir tableau 3). L'ajustement par le modèle gaussien simple est correct pour environ 99,3 % des ondes traitées. Une onde est considérée comme bien ajustée si le facteur de qualité $\xi < 0,5$. Il a été observé que $\xi < 0,1$ sur les zones urbaines avec un seul écho de forme gaussienne et que $\xi \leq 0,5$ même pour les ondes de forme complexe bien ajustées.

Utiliser une détection fine permet de trouver environ 5 % de points en plus qu'une simple détection et améliore l'ajustement des ondes (figure 6a). Comme attendu, le traitement d'ondes lidar a posteriori permet de pénétrer plus en profondeur dans les zones de forêts. Néanmoins, on ne peut toujours pas certifier que le dernier écho de l'onde correspond à une réflexion du sol sous la canopée. La figure 5 montre la différence d'altitude du dernier écho détecté entre les ondes traitées dans cet article et la méthode employée en temps-réel dans le vecteur aérien. Les points laser détectés a posteriori sont, de manière significative, plus proches du sol en zones forestières (jusqu'à 10 m). La détection des premiers échos de l'onde est également améliorée. Ces résultats sont récapitulés dans le tableau 3 où la différence moyenne ΔH entre l'altitude des échos détectés a posteriori et en temps-réel est toujours positive pour le dernier écho (ex. +1,58 m sur de la végétation dense) et toujours négative pour le premier écho (ex. -0,42 m).

Comparaison entre les modèles utilisés.

L'utilisation de la fonction **log-normale** améliore l'ajustement des échos asymétriques observables dans les rues et sur certains toits de bâtiments. Cependant, les ondes lidar ne peuvent pas être décomposées en une somme de fonctions log-normales car seuls quelques échos de chaque onde sont asymétriques.

Le mélange de modèles apparaît donc comme une perspective intéressante pour mieux traiter les données lidar.

La fonction **gaussienne généralisée** permet de modéliser à la fois les cas classiques d'échos gaussiens mais également ceux aplatis ou pointus. La figure 6b) montre que l'histogramme des valeurs de ξ pour le modèle gaussien généralisé est bien meilleur que le modèle gaussien. Néanmoins, un fort taux d'ajustements incorrects est noté (envi-

ron 0,4 % contre $< 0,1$ % avec la gaussienne). En théorie, la gaussienne généralisée devrait toujours estimer les ondes au moins aussi bien que la fonction gaussienne. Mais, en pratique, ce taux de divergence est dû à un problème de minimisation dans l'algorithme par moindres carrés. Dans tous les cas, le modèle gaussien généralisé n'est pas valable pour les échos asymétriques.

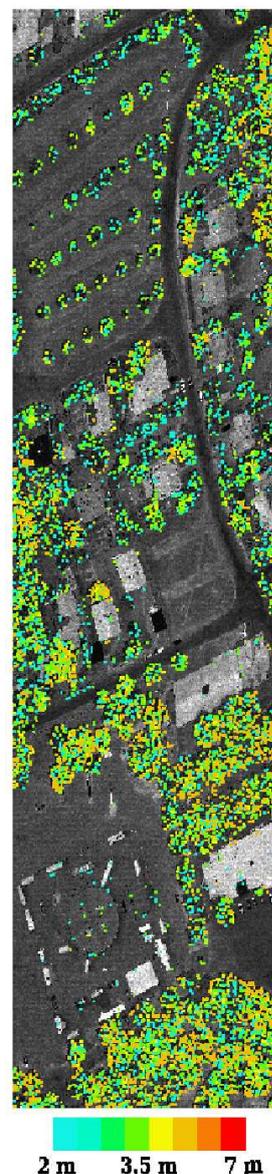


FIG. 5 – Différence entre les altitudes du dernier écho détecté par post-traitement et en temps-réel. Seules les différences supérieures à 2 m ont été affichées. L'image de fond est une image d'amplitude de premiers échos.

Comportement des paramètres des échos.

Une interprétation simple des paramètres associés aux échos extraits des ondes lidar (intensité, largeur et aplatissement) n'est pas possible et ne permet pas de définir une signature caractéristique pour chaque objet. Des para-

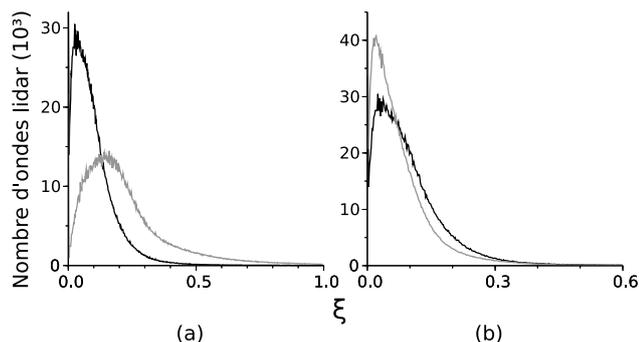


FIG. 6 – Histogramme des valeurs de ξ : **(a)** pour le modèle gaussien : méthode de détection basique (courbe noire) et fine (courbe grise); **(b)** pour un ajustement des ondes avec le modèle gaussien (courbe noire) et le modèle gaussien généralisé (courbe grise), avec une détection fine des échos présents dans l'onde.

mètres de cibles différentes peuvent avoir la même valeur sans toutefois que ces dernières aient des propriétés géométriques et radiométriques semblables. Ainsi, les routes et les toits de bâtiments ont des matériaux variables et donc des signatures variables, parfois similaires à des objets naturels.

Le paramètre supplémentaire de forme α obtenu par une modélisation plus générale des ondes ne permet pas non plus à lui seul la classification automatique des points laser. Même si aucune interprétation directe n'est donc possible, α s'avère moins sensible aux variations radiométriques des surfaces que les deux autres (présence de passage piétons, humidité des surfaces naturelles). Finalement, seul un processus englobant tous ces paramètres peut permettre une reconnaissance de formes dans une image via l'analyse de la morphologie des ondes. Sans aucune information supplémentaire sur ces paramètres (valeurs prises, corrélation entre eux), une classification supervisée par SVM est bien pertinente.

5.2 Résultats et évaluation de la classification

L'intégralité de la zone de test a été classée par SVM. La segmentation initiale comporte près de 12000 régions et environ 10 % des segments ont servi pour l'apprentissage de la classification. 7 % des régions initiales ont également été utilisées pour l'étape de validation croisée afin de trouver les valeurs optimales de C et γ . Le taux de classification correcte avec ces paramètres optimaux s'élève à 90,1 %. Cela illustre à la fois une certaine adéquation entre la base d'apprentissage et le problème d'apprentissage et l'absence de sur-apprentissage (le classifieur sait généraliser).

Les résultats de la classification sont visuellement satisfaisants : la distinction bâtiments–végétation haute–sol est bien réalisée. La figure 7 montre des résultats de classifica-

%	Bât.	Vég.	Sol nat.	Sol art.
Bât.	91,3	0,7	8	0
Vég.	7,5	92,5	0	0
Sol nat.	6,8	2,1	91,1	0
Sol art.	2,8	0	63,6	33,6

TAB. 4 – Matrice de confusion calculée à partir de la vérité terrain couvrant ~ 25 % de la zone de test ($\rho=0,77$).

tion sur le centre-ville de la zone test.

La séparation entre sol artificiel et sol naturel conduit à un taux d'erreur important. Cela est dû au fait que la zone de test possède en réalité peu de régions de sol naturel. Toutes ne pouvaient être intégrées dans la phase d'apprentissage pour ne pas fausser les résultats sur cette classe. Le classifieur n'a donc pu bien apprendre ses caractéristiques. L'étiquetage manuel est d'autant plus important que ces deux classes sont celles dont les attributs discriminants sont les plus proches (cf. tableau 2). En effet, si les zones de sols naturel et artificiel saisies lors de la phase d'apprentissage ont des réponses proches, en terme d'amplitude et de déformation du premier écho de l'onde retour, la classification devient difficile.

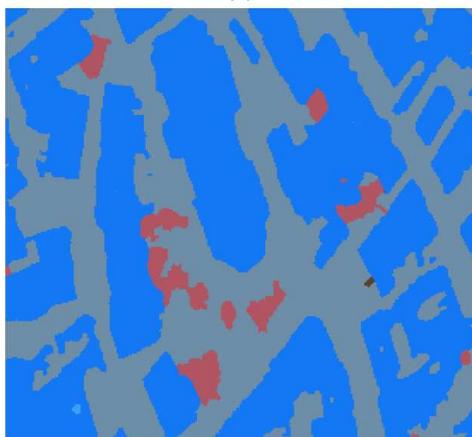
Le tableau 4 donne la matrice de confusion de la classification par SVM sur environ 2800 régions classées manuellement. On constate la faible différenciation des deux classes de sol, alors que les classes de bâtiments et de végétation sont bien segmentées. Comme attendu, certains bâtiments sont classés en sol artificiel (les valeurs des attributs peuvent se rejoindre) et inversement. Certaines zones de végétation sont étiquetées "bâtiment" quand leur taille est trop petite et que la variation d'altitude y est faible (cas d'arbres à houppier dense).

La taille des régions ont une influence non négligeable sur le résultat de la classification. Le choix de petites segments (proches d'une segmentation par watershed) assure une homogénéité des canaux (amplitude, altitude, α). Cependant, les nouveaux attributs introduits dans le SVM (intensité lidar, entropie, ...) ne sont pas forcément homogènes dans ces régions. Les petites régions sont alors plus sensibles à cette inhomogénéité que les grandes et la qualité de leur classification est en conséquence plus faible.

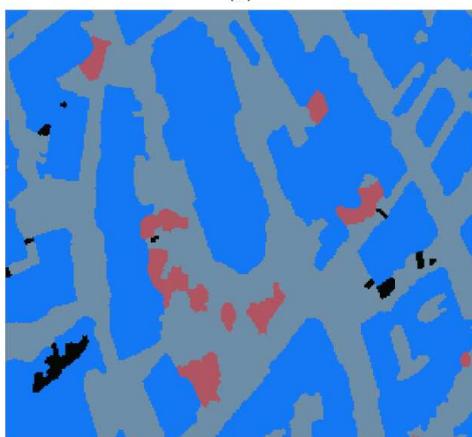
Les attributs du vecteur discriminant en entrée du classifieur SVM ne s'avèrent finalement pas tous pertinents. Pour évaluer leur apport, nous avons classé la zone de test tout d'abord avec uniquement le vecteur $\{\mu_{\Delta z}, \lambda_z\}$ puis ajouté progressivement les attributs discriminants (selon leur ordre d'apparition "historique", les systèmes laser fournissant tout d'abord une mesure d'altitude –entropie– puis l'intensité des échos et enfin leur largeur et leur forme). La figure 8 montre l'évolution du facteur de précision ρ selon les paramètres utilisés. L'ajout du terme d'entropie conduit à un plus haut taux de mauvaise classification ($\rho = 0,71 \rightarrow 0,69$). Cela est dû à certaines su-



(a)



(b)



(c)

FIG. 7 – Classification en milieu urbain : segmentation en classes "bâtiment" (bleu), "végétation haute" (rouge) et "sol artificiel" (gris). (a) Visualisation de la zone de test (ortho-image RGB), (b) Résultat en sortie de SVM, (c) Résultat affiné. Les régions en noir sont celles dont la probabilité d'appartenance à la classe choisie par le SVM est jugée trop faible (fixée à 60 %). Les images de labels sont affichées dans la géométrie du capteur laser.

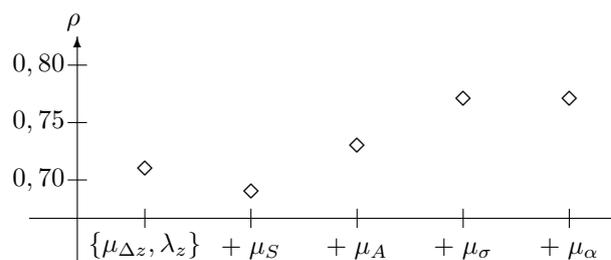


FIG. 8 – Évolution de ρ selon l'ajout de paramètres dans la classification par SVM. On part d'un vecteur d'attributs composé de $\mu_{\Delta z}$ et λ_z , puis on ajoute progressivement les attributs.

perstructures de toits (entropie forte) qui déclassent les régions "bâtiments" en "végétation haute" même s'il permet la séparation des zones de sol et de bâtiment. Les termes d'intensité et de largeur du premier écho des ondes lidar améliorent la segmentation en zone urbaine ($\rho = 0,69 \rightarrow 0,73 \rightarrow 0,77$) alors que l'ajout du terme α ne change pas la valeur de ρ . La raison semble être que l'image initiale sur-segmentée comprend un canal α , permettant la sur-segmentation de celles-ci en régions "homogènes" en α . La contribution de ce paramètre se fait à ce niveau-là et sa valeur moyenne par région n'apporte aucune information nouvelle.

Finalement, les termes μ_S et μ_α n'ont aucun apport pratique à l'algorithme de classification, malgré leur pertinence théorique. Cela souligne la difficulté de propager des attributs associés à des points laser à des régions. Rien en désigne en effet la moyenne comme la mesure statistique permettant la meilleure prise en compte des paramètres des points laser à l'échelle d'une région.

Par conséquent, il apparaît nécessaire de choisir de nouveaux attributs discriminants pour approfondir la classification par SVM. L'utilisation de la radiométrie des images semble permettre d'une part une meilleure segmentation des zones d'occupation des sols et d'autre part la différenciation de sous-classes dans les classes déjà bien segmentées (espèces d'arbres, matériaux de toits de bâtiments).

6 Conclusion et perspectives

Nous avons abordé dans cet article le problème de l'analyse et du traitement d'ondes lidar dans un objectif de reconnaissance de formes et de cartographie appliqué au milieu urbain. L'enregistrement de l'intégralité des signaux rétrodiffusés par la surface terrestre permet leur traitement a posteriori et améliore la connaissance des cibles atteintes par le faisceau laser.

Tout d'abord, une détection fine des échos contribuant à la déformation de l'onde est désormais possible. Elle conduit à la densification du nuage de points 3D final et donc à une meilleure connaissance de la distribution altimétrique des objets rétrodiffusants sur la surface terrestre. D'autre part, une fois la décomposition des ondes réalisée, nous

avons proposé leur modélisation par différentes fonctions afin de conserver pour chacun des échos ses caractéristiques de forme. Les fonctions log-normale et gaussienne généralisée se sont avérées toutes les deux pertinentes et contribuent à l'ajustement de pics auparavant mal estimés par une simple fonction gaussienne. Ces deux fonctions ont par ailleurs leurs contributions spécifiques c'est pourquoi elles doivent être toutes les deux conservées. D'autres fonctions avec d'autres paramètres doivent également être utilisées afin de tester si elles peuvent apporter une connaissance supplémentaire sur les formes des échos des ondes. Nous envisageons pour cela d'implémenter un algorithme de type Monte Carlo par Chaînes de Markov à sauts réversibles (RJCMC) afin d'extraire pour chaque onde le meilleur mélange de fonctions et d'utiliser des fonctions non optimisables par simples Moindres Carrés.

La modélisation des ondes lidar apporte pour chaque point des attributs supplémentaires à la simple géométrie 3D. Un traitement possible de ces résultats est la segmentation des données lidar. Nous avons par conséquent proposé une méthode de classification supervisée par SVM. Les résultats sont encourageants et montrent le potentiel d'une segmentation des zones urbaines en classes bâtiments-végétation haute-sol naturel-sol artificiel. Elle produit des masques de focalisation corrects pour des traitements futurs. Elle doit toutefois être approfondie pour conduire à de meilleurs résultats.

La première étape est la recherche un noyau dédié à la classification de données lidar. Cela passe forcément par la connaissance de l'influence respective des attributs des échos extraits. Nous envisageons finalement d'affiner la méthode de classification via un processus itératif en se servant d'images aériennes orientées et de nouveaux paramètres extraits des ondes. Un premier pas vers la reconnaissance de formes en milieu urbain est réalisé avec l'algorithme proposé. L'apport de la radiométrie des images couplée aux paramètres d'ondes lidar nous permet ensuite d'envisager une classification, à résolution plus fine, des espèces d'arbres, des matériaux de toits et une séparation des terrains artificiels et naturels.

Références

- [1] J. Kilian, N. Haala, M. Englich. Capture and evaluation of airborne laser scanner data. *IAPRS, Vol. XXXI*, Vienne, Autriche, 1996, pp. 383–388.
- [2] J. Blair, D. Rabine, M. Hofton. The Laser Vegetation Imaging Sensor : a medium-altitude digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (1999) 115–122.
- [3] R. Dubayah, J. Blair. Lidar Remote Sensing for Forestry Applications. *Journal of Forestry* 98 (6) (2000) 44–46.
- [4] H. Zwally *et al.* ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics* 34 (3) (2002) 405–445.
- [5] A. Persson, U. Söderman, J. Töpel, S. Alhberg. Visualization and Analysis of Full-Waveform Airborne Laser Scanner Data, *ISPRS workshop 'Laserscanning 2005'*, Enschede, Pays-Bas, 2005, pp. 103–108.
- [6] M. Hofton, J. Minster, J. Blair. Decomposition of Laser Altimeter Waveforms. *IEEE Trans. on Geoscience and Remote Sensing* 38 (4) (2000) 1989–1996.
- [7] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, N. Studnicka. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (2006) 100–112.
- [8] B. Jutzi, U. Stilla. Laser pulse analysis for reconstruction and classification of urban objects. *Photogrammetric Image Analysis (PIA)*, Munich, Allemagne, 2003, pp. 151–156.
- [9] H. Gross and B. Jutzi and U. Thoennessen. Segmentation of Tree Regions using Data of a Full-waveform Laser. *Photogrammetric Image Analysis (PIE)*, Munich, Allemagne, 2007, pp. 57–62.
- [10] K. Thiel, A. Wehr. Performance and capabilities of laser scanners - an overview and measurement principle analysis. *ISPRS 'Laser-Scanners for Forest and Landscape Assessment'*, Fribourg, Allemagne, 2004.
- [11] O. Steinvall. Waveform simulation for 3-D sensing laser radars. *Rapport Technique du FOI* (Mai 2000).
- [12] W. Wagner, A. Ullrich, T. Melzer, C. Briese, K. Kraus. From single-pulse to full-waveform airborne laser scanners : Potential and practical challenges. *IAPRS* (2004) 201–206.
- [13] V. Ducic, M. Hollaus, A. Ullrich, W. Wagner, T. Melzer. 3D Vegetation mapping and classification using full-waveform laser scanning. *Workshop on 3D Remote Sensing in Forestry*, Vienne, Autriche, 2006.
- [14] C. Bishop. Pattern Recognition and Machine Learning, *Springer, New-York*, 2006.
- [15] B. Aiazzi, L. Alparone, S. Baronti. Estimation based on entropy matching for generalized Gaussian PDF modelling. *IEEE Signal Processing Letters* 36(6) (1999) 138–140.
- [16] L. Guigues, J.P. Cocquerez, H. Le Men. Scale Sets Image Analysis. *International Journal of Computer Vision* 68(3) (2006) 289–317.
- [17] L. Guigues, Modèles Multi-Échelles pour la Segmentation d'Images, Thèse de doctorat de l'Université de Cergy-Pontoise, 2004.
- [18] J. Secord, A. Zakhor. Tree Detection in Urban Regions Using Aerial LiDAR and Image Data. *IEEE Conference on Image Processing*, Atlanta, États-Unis, 2006.
- [19] S. Canu. Apprentissage et noyaux : séparateur à vaste marge (SVM). *Revue de l'électricité et de l'électronique* 8 (2006) 69–78.
- [20] C.-W. Hsu, C.-J. Lin. A comparison of methods for multi-class support vector machines. *IEEE Trans. on Neural Networks*, 13 (2002) 415–425.
- [21] M. Pontil, A. Verri. Properties of support vector machines, *Rapport de Recherche du MIT*, 1998.