Photogrammetrie • Fernerkundung • Geoinformation 5/2008, S. 337–349, 6 figs., 3 tab.

Analysis of Full-Waveform Lidar Data for Classification of Urban Areas

Clément Mallet, Frédéric Bretar, Saint-Mandé/Frankreich & Uwe Soergel, Hannover

Keywords: Lidar, signal processing, waveform analysis, modelling, classification, urban

Summary: In contrast to conventional airborne multi-echo laser scanner systems, full-waveform lidar systems are able to record the signal of each emitted and backscattered signal of each laser pulse. Instead of clouds of individual 3D points, lidar devices provide connected 1D profiles of the 3D scene, which contain more detailed and additional information about the structure of the illuminated surfaces. This paper is focused on the analysis of full-waveform data in urban areas. First, a brief overview of the workflow of the proposed classification model is given. Then the problem of modelling full-waveform lidar signals is tackled. The standard method assumes the waveform to be the superposition of signal contributions of each scattering object in such a laser beam, which are approximated by Gaussian distributions. This model is suitable in many cases, especially in vegetated terrain. However, since it is not tailored to urban waveforms, here the generalized Gaussian model is selected instead. Then, a pattern recognition method for urban area classification is proposed. A supervised method using Support Vector Machines is performed on the full-waveform lidar point cloud based on the parameters extracted from the post-processing step. Results show that it is possible to partition urban areas in building, vegetation, natural ground and artificial ground regions with high accuracy using only lidar waveforms.

Zusammenfassung: *Analyse des zeitlichen Signalverlaufs von Laserpulsen zur Klassifikation städtischer Gebiete*. Daten konventioneller luftgestützter Laserscannersysteme beschränken sich meist auf die Koordinaten der 3D-Punkte, gegebenenfalls erweitert durch weitere Merkmale, wie etwa Breite und Intensität der Echos. Eine neue Generation von Sensoren erfasst durch Abtastung den zeitlichen Verlauf der Sendepulse sowie des empfangenen Signals. Im Gegensatz zu den bislang verfügbaren diskreten Punktwolken entspricht die mit diesen sogenannten Full-Waveform-Sensoren gewonnene Signalform somit zusammenhängenden 1D-Profilen der 3D-Szene, anhand derer weitergehende Rückschlüsse auf die Struktur der beleuchteten Oberflächen möglich sind. Der Schwerpunkt dieses Aufsatzes liegt auf der Auswertung von Full-Waveform-Daten, die über städtischem Gebiet gewonnen wurden. Nach einer Übersicht des vorgestellten Verfahrens zur Klassifikation solcher Profile, wird das zugrunde liegende Datenmodell erläutert. Im Standardmodell ergibt sich die Signalform aus einer Überlagerung der Beiträge aller sich im Kegel des Laserstrahls befindlichen Streukörper, die in guter Näherung mit Gaußkurven approximiert werden können. Dieses Modell ist in vielen Fällen angemessen, insbesondere bei Vegetationsflächen. Zur Beschreibung urbaner Gebiete ist es allerdings nicht hinreichend, daher wird hierzu in diesem Beitrag ein erweitertes generalisiertes Gaußmodell vorgeschlagen. Zur überwachten Klassifikation der Daten dient eine Support-Vector-Maschine, die auf in einem Vorverarbeitungsschritt aus den Profilen extrahierten Merkmalen aufsetzt. Die Ergebnisse zeigen, dass es möglich ist, anhand der Signalformen mit hoher Genauigkeit die Klassen Gebäude, Vegetation, natürliches Gelände und künstliches Gelände zu unterscheiden.

1432-8364/08/2008/337 $ 3.25

1 Introduction

In the last decade, airborne lidar systems have become an alternative source for acquisition of altimeter data. Such devices deliver a reliable, fast and accurate representation of terrestrial landscapes through georeferenced and unstructured 3D point clouds (root mean square error, RMSE < 0.1 m in altimetry). Range is determined directly according to the signal runtime measurement, whereas stereoscopic techniques derive the 3D information, indirectly based on the camera orientations and the disparity of correspondences in optical images identified by image matching methods. A large body of literature addresses the potential of laser scanning data for urban and suburban area analysis. For instance, many algorithms for classifying lidar point clouds have been developed so far aiming at building detection and subsequent reconstruction (Haala & Brenner 1999, Sithole & Vosselman 2004).

Often external context information, such as cadastral maps, is exploited. However, even in the case of lack of supplementary data, generally at least the main building outlines can be extracted from the laser point clouds alone. In the latter case the discrimination of buildings from adjacent trees is difficult. All these approaches rely only upon geometric and topologic criteria and have in common that they are sensitive to large off-terrain objects and surface discontinuities. Therefore, many authors proposed to add other inputs like echo intensity (Tóvári & Vögtle 2004) or multi-spectral images (Rottensteiner et al. 2005), which are often gathered in parallel to the laser data from the same carrier today, to achieve better results.

Since a few years, a new generation of lidar devices designed to digitize and to record the entire signal of each emitted laser pulse became operational. They are called **full-waveform** lidar systems. Full-waveform data offer the opportunity to overcome many drawbacks of classical multi-echo lidar data (Wagner et al. 2004). In addition to single range measurements, further physical properties of the objects included in the diffraction cone may be revealed by analysis of the shape of the sampled backscatter sequence.

Many studies have already been carried out to perform full-waveform data processing, mainly in vegetated areas. The higher point density inside the penetrated canopy offers insight in the vegetation type and state (Harding et al. 2001). In urban areas, the potential of such data has been barely investigated. For instance, Jutzi & Stilla (2005) extract linear features on roof tops by establishing neighbourhood relationships between waveforms. They also localize more accurate building outlines. Moreover, Kirchhof et al. (2008) iteratively process terrestrial waveforms to detect additional points in partially occluded and partly illuminated regions. On the other hand, by exploiting other features in addition to the geometry (*e. g.*, pulse amplitude or width), detection of vegetated areas is now possible (Gross et al. 2007). To achieve more advanced point classification in urban areas, a theoretical knowledge of the influence of the geometric and radiometric properties of the hit targets (i. e., the differential laser cross-section) on the shape of the lidar waveforms is required.

The aim of the article is to show that a fine analysis of full-waveform lidar data can lead to an accurate classification of urban areas. In Section 2, the general outline of this work is described. Then, a new modelling function is proposed in Section 3 to process raw signals. Results of the integration of the previously extracted features into a supervised classification algorithm are presented in Section 4. The aim is to discriminate four classes: buildings, vegetation, artificial ground, and natural ground regions. The test data sets are outlined in Section 5. Finally, the results of waveform processing and classification are presented and the conclusions are drawn.

2 Overall Methodology

Common laser data formats are clouds of 3D points, often provided without link to the original laser shot. In contrast to this,

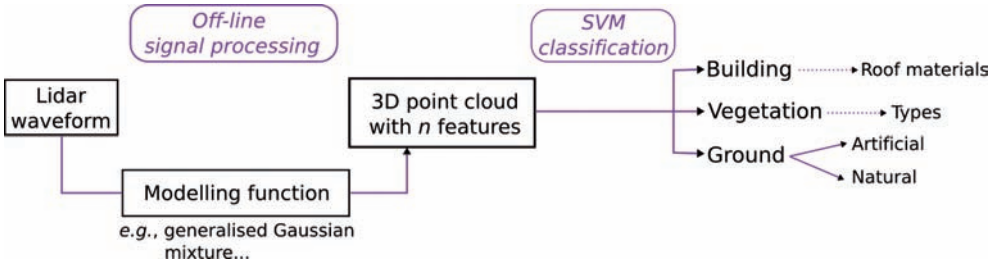

Fig. 1: Lidar waveform processing workflow. Echoes within each waveform are first modelled using an analytical function. The generated 3D point cloud is then classified using features extracted from the signal processing step.

lidar profiles comprise information of the 1D object structure along the line of sight. Nevertheless, such data are more difficult to handle and specific studies have to be carried out. In this article, an approach is proposed to process full-waveform lidar data to extract 3D point clouds featuring more useful parameters in addition to the traditional (x, y, z) coordinates and to subsequently perform a point classification based on these parameters (see Fig. 1).

Waveform processing consists in decomposing the waveform into a sum of components or echoes, in order to characterise the different individual targets along the path of the laser beam. A parametric approach is chosen, i. e., parameters of a mathematical model are estimated. Non-analytic methods like neural networks, or Parzen windows are known to work very well to approximate 1D signals (Bishop 2006). Nevertheless, they do not provide physical information about the target (laser cross-section). The objective of waveform processing is twofold. It starts in maximizing the detection rate of relevant peaks within each signal in order to foster information extraction from the raw signal. Furthermore, from a class of functions the best fit to the waveform is chosen. This allows to introduce new parameters for each echo and to extract additional information about the target shape and its reflectance.

Then, the extracted point cloud is classified. The aim of our work is to assess whether or not each new parameter introduced is a relevant feature for classification and how significant it is for urban analysis (does it provide useful information?). The features are fed into a supervised classification algorithm using a Support Vector Machine (SVM). This method is well adapted to deal with high-dimensional feature space since the algorithm complexity does not depend on the data dimension. Furthermore, SVM belong to the non-parametric classification techniques, i. e., no parametric probability density functions are required. In recent years, SVM became relevant for solving remote sensing classification tasks. SVM permits to use jointly classical geometric features, image-based information (Secord & Zakhor 2007) as well as in our case new parameters extracted from the post-processing step.

The methodology for classifying urban areas by lidar waveform analysis is designed to be flexible. Depending on the modelling function, the theoretical understanding of pulse propagation in such regions and the chosen options of the SVM classifier, it is possible to refine the classification (*e. g.*, by splitting the "vegetation" class in tree species or "buildings" in roof materials).

3 Waveform Processing

Our methodology is based on a paper written by Chauve et al. (2007). The authors describe a waveform processing using an iterative Non-Linear Least Squares fitting algorithm. After coarse initial peak detection, missing peaks are found in the difference between the modelled and the initial signal. If new peaks are detected, the fit is performed

again. This process is repeated until no further improvement is possible. This enhanced peak detection method is useful to model complex waveforms with overlapping echoes and also to extract weak echoes not found by on-line detection techniques. In urban areas, the former case is observed when the laser beam hits building edges. The resulting waveform is therefore composed of distributed backscatters of the roof and the ground, which can often not be separated by hardware detection algorithms using thresholds.

Waveforms collected with a small-footprint lidar system are used in this article (RIEGL LMS-Q560). Such data can be well modelled by superposition of Gaussian pulses. Wagner et al. (2006) have shown that more than 98 % of the observed waveforms collected from the RIEGL system could be approximated by a sum of Gaussian functions. Each laser output pulse shape is assumed to be Gaussian, with a specific and calibrated width. The collected pulse is therefore a convolution between this distribution and a "surface" function, depending on the reflecting objects. Nevertheless, the transmitted signal is not always Gaussian. Indeed, it is observed that the LMS-Q560 transmitted waveform is slightly asymmetric.

In urban areas, most of the return waveforms are in reality subject to the mixed effects of geometric (*e. g.*, roof slopes) and radiometric object properties (*e. g.*, different kinds of streets and roof materials). Histograms of the four considered classes are illustrated in Fig. 2. Hence, the characteristics of return peaks may differ significantly. It was already shown that standard extensions of the Gaussians model, such as Lognormal and generalized Gaussian functions, are suitable to model raw lidar signals. Using the generalized Gaussian (GG) model improves the signal fitting for symmetric and distorted waveform shapes (more than 99.3 % of satisfactory results) (Chauve et al. 2007). The authors argue that the Lognormal model fits asymmetric pulses with success but fails for symmetric ones.

Here, the GG model was used to process full-waveform lidar data:

$$f_{GG} = A \cdot \exp\left(\frac{|x-\mu|^{\alpha^2}}{2\sigma^2}\right) \quad (1)$$

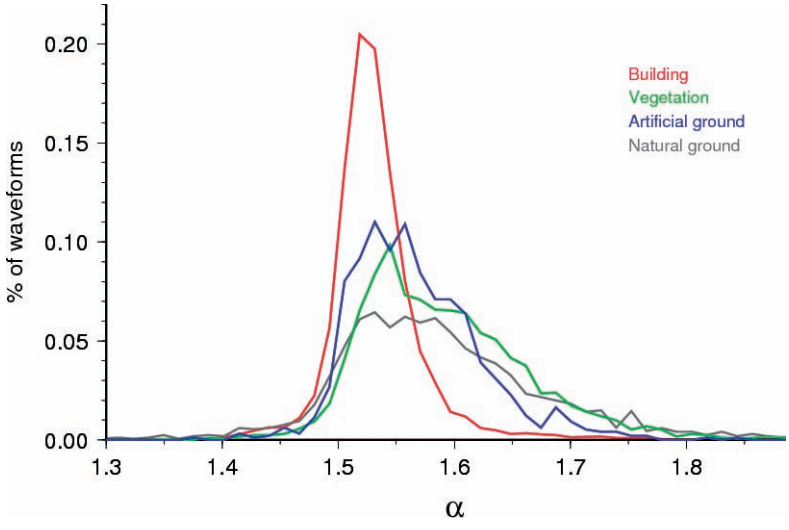

Fig. 2: Histogram of α values over the four test classes. Each plot indicates, for the respective class, the percentage of echoes having a specific α value included in the range [1.3, 1.9]. α is one of the parameters extracted from the signal processing step based on a generalized Gaussian modelling of the peaks. α = 1.41 simulates the Gaussian function, typically used for aerial lidar waveform modelling.

where A is the pulse amplitude, σ its width, μ the function mode and α the shape parameter which allows to simulate Gaussian ($\alpha = \sqrt{2}$), flattened ($\alpha > \sqrt{2}$) or peaked ($\alpha < \sqrt{2}$) pulses.

A, σ and α are the three new features which will be introduced in the classification step in Section 4.

4 Classification of Urban Areas

4.1 *Methodology*

Based on a clustering analysis of the parameters extracted from the modelling step, four classes have been chosen to characterize urban areas: buildings, vegetation, natural ground, and artificial ground. Artificial ground gathers all kinds of streets (tar, asphalt...), pavements and, street items (*e.g.*, cars or traffic lights) whereas the natural ground class includes grass, sand, and bare-earth regions.

4.2 *Support Vector Machines*

The general mathematical formulation of SVMs is briefly recalled in this section.

Linear SVMs

Let D be the data space, Y the label space and A the training set (*e.g.*, $D = \Re^n$, $Y = \{-1, 1\}$ in a two-class problem).

The classification is carried out using a linear discriminant function $\omega(D \to Y)$. $x_i \in A$ are the N training samples available with their labels $y_i / i \in [1, N]$. The theoretical aim of supervised classification is to find a classifier consistent with the training set. The SVM method consists in finding the hyperplane maximizing the distance (called the margin) to the closest training data points in both classes (the support vectors). For a linear classifier, $\omega(x) = \mathbf{w}.x\text{-}\theta$, where $\mathbf{w} \in D$ is the normal vector to the hyperplane and θ the bias. We aim at finding the classifier parameters (w, θ) which verify:

$$\forall (x_i, y_i) \in A,\ y_i \times (w \cdot x_i - \theta) > 0 \qquad (2)$$

Since the SVM method searches the best classifier (*i.e.*, the largest margin), we impose:

$$\forall (x_i, y_i) \in A,\ y_i \times (w \cdot x_i - \theta) \geq 1 \qquad (3)$$

The support vectors lie on two hyperplanes $w \cdot x_i - \theta = \pm 1$ which are parallel and equidistant to the optimal linear separable hyperplane. Finally, the optimal hyperplane has to maximize the margin (*i.e.*, the Euclidian distance between both hyperplanes, defined as $2/\|w\|$ under the constraints defined in Equation 3). Unfortunately, in most cases, such quadratic optimization problem is unsolvable: we cannot find a linear classifier consistent with the training set, because the classification problem is not linearly separable.

Consequently, slack variables[1] ξ_i are introduced to cope with misclassified samples and prevent Equation 3 from being violated. Another reason is the avoidance of over-fitting the classifier to the training samples, which would result in poor performance. Equation 3 thus becomes:

$$\forall (x_i, y_i) \in A,$$
$$y_i \times (w \cdot x_i - \theta) > 1 - \xi_i / \forall_i \in [1, N],\ \xi_i \geq 0 \qquad (4)$$

C is a constant which determines the trade-off between margin maximization and training error minimization.

The final optimization problem is subsequently:

$$\min \left[\frac{\|w\|^2}{2} + C \sum_{i=1}^{N} \xi_i \right] \quad \text{subject to (4)} \qquad (5)$$

Nonlinear SVMs

When the classification problem is not linearly separable, one solution consists in changing the feature space. The data is projected in a higher dimension space using a nonlinear mapping function Φ: $D \to H$, in which the new distribution of samples enables the fitting of a linear hyperplane. Ker-

[1] A slack variable is a nonnegative variable that turns an inequality into an equality constraint.

nel methods provide nonlinear hyperplanes and improve classification abilities. The same margin optimization method can then be performed.

Finding Φ is a difficult problem. In practise, the xi points are implicitly projected into H by defining a kernel K: $D \times D \rightarrow \Re$ with $K(x_i, x_j) = (\Phi(x_j)|\Phi(x_j))$. In fact, the knowledge of K is sufficient to compute the optimal classifier. It has only to fullfil Mercer's condition (Schölkopf et al. 1998).

Multiclass SVMs

SVMs are designed to solve binary problems. When having $n > 2$ classes of interest, various approaches are possible to address the problem, usually combining a set of binary classifiers. We choose the 'one-against-one' approach because it has been shown to be more suitable for large problems (Hsu & Lin 2002). For such pairwise classification, $n(n-1)/2$ binary classifiers are computed on each pair of classes. Each sample is assigned to the class getting the highest number of votes. A vote for a given class is defined as a classifier assigning the sample to that class.

In practise

The LIBSVM software is used to implement the SVM algorithm (available at www.csie.ntu.edu.tw/~cjlin/libsvm). Slack variables are introduced (soft-margin classifier). Then, the parameter C has to be optimized with the kernel hyperparameters (see Section 4.3).

4.3 *Kernel Selection*

Without sufficient a priori knowledge of the influence of geometric and radiometric target properties on the pulse shape (or even strong hints about characteristic behaviours on urban areas), it is difficult to design a kernel dedicated to our purpose. Therefore, the generic Gaussian kernel was selected. It is defined as:

$$K(x_i, x_j) = \exp\left(\frac{\| x_i - x_j \|^2}{2\gamma^2}\right) \quad (6)$$

where $x_{i,j}$ is the data to be classified and γ is used to express how similar to the training data the test data is expected to be ($\gamma \rightarrow 0$ for instance leads to over-fitting and consequently reveals a low generalization ability of the classifier). Because optimal values of C and γ are not known beforehand, a grid search is performed in which the cross-validation accuracy (CVA) is computed for each point (the CVA over (C, γ) set is not convex). A v-fold cross-validation procedure is carried out *i.e.*, the training data are divided in v subsets of equal size. The classifier is trained on v-1 subsets and ran on the remaining one. Such process has the advantage of not requiring a separate and independent data set for assessing the classification accuracy. The CVA represents the percentage of samples correctly classified averaged over all the subsets when they were used as the testing subset. After the coarse grid search, a finer one is computed in a smaller range around the optimal parameters found in the first step. Grid search is not necessary but is one simple tool for identifying the optimal hyperparameters. Several advanced methods exist. But, for two parameters, the computational time to find (C, γ) with such exhaustive search is not significantly higher than genetics or gradient-based optimization algorithms (Chapelle et al. 2002).

4.4 *Feature Selection and Relevance*

Our feature vector for each lidar point has **eight** components.

- $\boldsymbol{\mu}_{\Delta r}$: difference between the pulse range and the highest range (lowest altitude) found in a large spherical environment (20 m radius for instance; this parameter is selected manually, depending on the data set),
- **R**: residuals computed from a plane estimated by a robust M-estimators with norm $L_{1.2}$ (the selection of an optimal p has been investigated and for p around 1.2, a good estimate may be expected, see (Xu & Zhang 1996) for more details) on the points in a given neighbourhood

Tab. 1: Empirical values of the selected features for SVM classification for the four labels.

Feature	Building	Vegetation	Artificial Ground	Natural Ground
$\mu_{\Delta r}$	variable	variable	$\rightarrow 0$	$\rightarrow 0$
R	$\rightarrow 0$	high	$\rightarrow 0$	$\rightarrow 0$
n_Z	$[-45°, 45°]$	variable	$[-10°, 10°]$	$[-10°, 10°]$
Δz_{fl}	0	high	0	0
N	≈ 1	≥ 1	≈ 1	1
A	variable	medium	low	variable
σ	medium	high	variable	variable
α	[1.5, 1.6]	variable	≈ 1.6	≥ 1.6

(here, a spherical environment of 0.5 m radius),
- $\mathbf{n_z}$: the deviation of the local normal vector from the vertical direction,
- $\mathbf{\Delta z_{fl}}$: the altitude difference between the first and the last pulse of the waveform,
- **N**: the number of echoes within the waveform of the current lidar point,
- **A, σ, α**: the pulse amplitude, width, and shape, respectively (extracted from the waveform processing step described in Section 3).

The first three parameters can be used with every 3D point cloud (only geometric information). The last three ones are derived by waveform modelling (amplitude can also be available with multiple-pulse point clouds). Feature $\mu_{\Delta r}$ allows to filter points on the terrain from off-terrain points; Δz_{fl} and N discriminate vegetation points from the others. Multiple reflections can occur when the laser beam hits a building (due to superstructures, *e. g.*, chimney, roof edges) or a street (due to objects on the street e. g., vehicles or traffic lights). R and n_z values are also affected in these cases. The generalized Gaussian parameters are introduced in the SVMs to see how significant they are for the segmentation between the four classes and in between natural and artificial grounds.

Tab. 1 summarizes the feature values for the different labels. Other features have been tested such as the altimetric texture and several moments of the three extracted parameters in a given neighbourhood (mean, standard deviation, and skewness) and the backscatter cross-section (Wagner et al. 2006) but they were not found relevant for our study.

5 Full-waveform Lidar Data

Two data sets are available for this study. Data acquisition was performed with the RIEGL LMS-Q560 system over the cities of Biberach (Germany) and Le Brusquet (France). The main technical characteristics of this sensor are presented in (Wagner et al. 2006). The specifications of the two surveys are described in Tab. 2.

Each return waveform is a signal composed of one or two sequences of 60 and 80 samples (for Biberach and Le Brusquet, respectively), *i. e.*, 60 and 80 bins with a recorded amplitude. For each recorded waveform, the digitized emitted pulse and the echoes found by the hardware detection algorithm are given as well as their amplitude and width. In urban areas, the digitization of vertical sections of 60 ns (around 18 m) is sufficient to record backscattered signals both from the tree tops and the ground below them.

The city of Biberach includes residential, industrial, and dense urban areas. The surveyed area of Le Brusquet consists of scattered houses in a rural region.

Tab. 2: Overview of the specification of the data sets (PRF = pulse repetition frequency).

Area	Urban specificity	Flight height (m)	Footprint size (m)	PRF (kHz)	Pulse width (ns)	Temporal sampling (ns)	Pulse density (/m²)
Biberach	dense	500	0.25	100	> 5	1	2.5
Le Brusquet	sparse	700	0.35	111	> 5	1	> 5

An artifact, specific to the RIEGL sensor and called 'ringing effect', is noticed within the waveforms: after the sampled emitted pulse, a small secondary maximum resulting from effects of the hardware waveform processing chain can be seen. Consequently, in urban areas, when the laser beam hits planar objects of high reflectance and with a small angle of incidence (typically streets and roofs), such artifact is still present in the reflected waveform. In the iterative process (see Section 3) a weak pulse just behind a strong one is therefore removed when their amplitude ratio is close to the ratio computed from the emitted pulse.

6 Results and Discussion

6.1 *Modelling Raw Signals*

As described in detail in (Chauve et al. 2007), it is still appropriate to model complex waveforms with the GG function and all the more crucial in urban areas. Indeed, the benefits of full-waveform data for building reconstruction or classification are twofold:

First, the GG model improves signal fitting. More relevant points are extracted. Compared to the standard method 5% additional pulses are found which correspond to weak pulses in trees, hedges, building edges and roof superstructures. Furthermore, taking the 'ringing effect' into account allows to exclude artifacts (*i. e.*, non-existing points) during post-processing (see Fig. 3). On ground and building regions, ringing points are removed (> 15% of the total number of points).

Secondly and above all, the global signal fitting quality is increased. Flattened and high single pulses as well as narrow ones are now well detected. Fig. 2 shows that since α values are in many cases larger than $\sqrt{2}$ (mean value = 1.52), waveforms are in reality flattened, compared to Gaussian curves. Depending on the application, the Gaussian model can nevertheless be sufficient. For

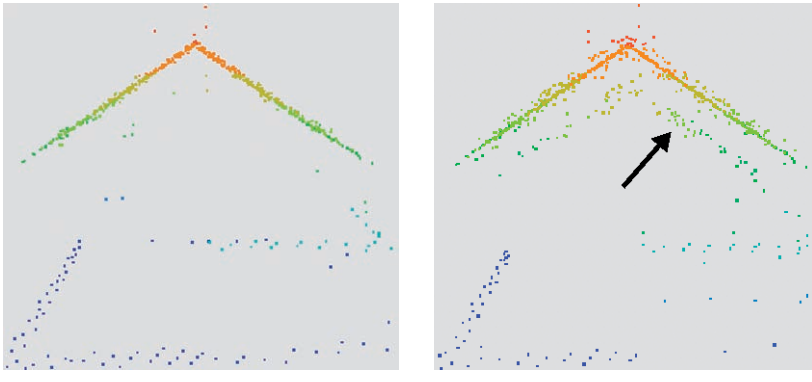

Fig. 3: Building point cloud without taking the 'ringing effect' into account (left, the black arrow shows the false point layer). The same data after the removal of the sensor artifacts (right). The roof no longer appears doubled.

example, in forested areas waveforms are mainly composed of at least two peaks. In such application, it is often not of interest to extract a shape parameter, which will depend both on the reflected target and on the targets already hit by the laser beam. But, in urban areas, the GG contribution is all the more significant since this parameter provides genuine information about the target shape and reflectance.

6.2 *Behaviour of Extracted Parameters*

A visual interpretation of the shape of lidar waveforms is needed and a simulation step is required to understand how the pulse interacts with the targets and to decorrelate geometric and radiometric influences. Amplitude and width values have also to be corrected according to the waveform angle of incidence and the target slope. Analysis of extracted point clouds revealed the following general behaviour of the three extracted parameters for different targets in urban areas:

- High **amplitude** values are found on building roofs independent of the material (except metal), on gravel, on sand, and cars. The lowest values correspond to vegetation points, due to a higher target heterogeneity and attenuation. Asphalt and tar streets have also low amplitude values, but despite low contrast it is possible to visually discriminate different kinds of surfaces.
- Vegetation spreads lidar pulses; that is why the highest **width** values are found in trees and hedges. Ground and building surfaces coincide with low width values, even if it is noticed that an increasing roof slope tends to increase pulse width.
- Very low and high **shape** values are characteristics of building edges and vegetation. Building regions correspond to α values in a specific range (between 1.5 and 1.6). Natural ground (especially grass) and artificial ground surfaces can also be visually distinguished. However, vegetated areas exhibit comparable values (see Fig. 2).

6.3 *Classification*

Both data sets have been classified. Approximatively 0.8 % of the pulses were used for the training step and 1 % to find the optimal values of C and γ.

Tab. 3 gives the classification results over the city of Biberach using the vector composed of eight features. It shows that the segmentation between different kinds of ground leads to a certain rate of misclassification. The main reasons are, first, that only few grass or sand regions are present in the Biberach area and therefore only limited numbers of samples are available for training and test. Moreover, the clusters in the feature space of these two classes are very close (*cf.* Tab. 1). The results are therefore very sensitive to the training step and the selected regions. Consequently, the SVM classification often fails when discriminating these two regions. Nevertheless, tests carried out on the city of Le Brusquet (rural area) show that classification in four labels is still

Tab. 3: Confusion matrix computed with ground truth consisting on 6 % of the whole data set of Biberach ($\rho = 0.81$ and 135627 points).

Number of points	% points correctly classified	Building	Vegetation	Artificial Ground	Natural Ground
76593	Building	**87.1**	8.8	3.6	0.5
8943	Vegetation	10.2	**88.9**	0.7	0.2
49048	Artificial Ground	2.2	2.1	**84.6**	11.1
1043	Natural Ground	4.1	~ 0	33.2	**62.7**

conceivable when enough training samples are available (*cf*. Fig. 5). The *building* and *vegetation* points are well classified. As expected, some building points are classified as *ground* (their values can be close, e. g., a flat dark roof close to the ground) and as *vegetation* especially superstructure and building edge points. Vegetated points can also be labelled as *building* when the laser beam hits dense tree areas.

The overall accuracy ρ is used as a quality criterion and is defined as:

$$\rho = \frac{\sum_{i=1}^{\dim Y} A_{i,j}}{\sum_{i=1}^{\dim Y} \sum_{j=1}^{\dim Y} A_{i,j}} \in [0,1] \qquad (7)$$

where $A_{i,j}$ gives the number of laser points labelled as *j* and belonging to the class *i* in reality. ρ is equal to 1 when the classification is perfect and 1/(number of labels) when the classifier randomly chooses the class for each point with the same probability. Fig. 4 shows the evolution of the classification accuracy depending on the input features, adding them by their historical 'order of appearance' (see part 4.4). The first five attributes are available or can be computed with multi-echo lidar data. Amplitude is sometimes given with the 3D point cloud that is why this feature is then added. The width and shape attributes are finally introduced, beginning with σ to assess the contribution of the generalized Gaussian waveform decomposition compared to the standard Gaussian one. Each new feature improves the classification results but such result strongly relies on the order of feature introduction.

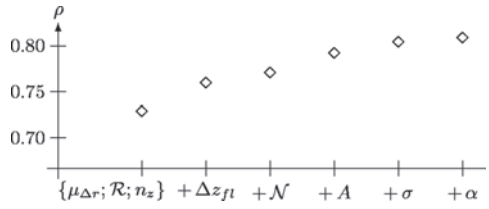

Fig. 4: Overall accuracy evolution depending on the features included in the SVM algorithm. Starting from the vector $\mu_{\Delta r}$, R, n_z, the other ones are added progressively (Biberach area).

A label-by-label analysis reveals first that the amplitude value allows to discriminate building and ground points: ground surfaces have much lower reflectance than roof materials. Moreover, although it is noticed that artificial ground surfaces have lower amplitude values than natural ones, the difference is barely significant and does not permit their discrimination. Feature σ is helpful to enhance the building/vegetation separation but not for ground region segmentation. Results are slightly worse for ground points with α than without the integration of this parameter for the Biberach data sets (63.3% success without α for the *natural ground* class), whereas this parameter visually improves the results over Le Brusquet (see Fig. 5, no ground truth available for this area). Finally, feature α slightly improves point cloud classification by better discriminating building points. Although natural

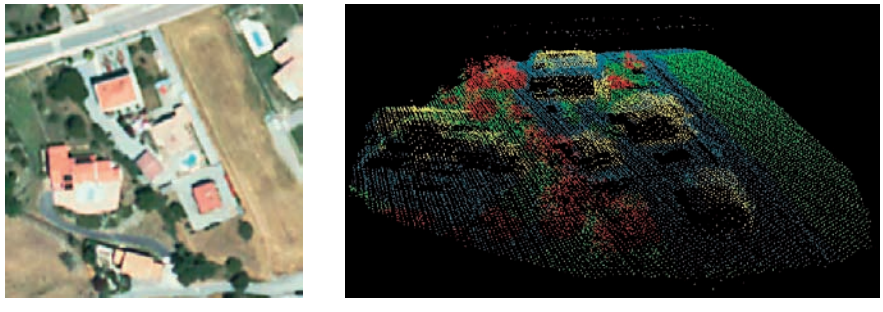

Fig. 5: Classification results in a scattered urban area (Le Brusquet). Left: orthoimage of the region of interest. Right: classified point cloud (yellow: buildings, red: vegetation, blue: artificial ground and green: natural ground).

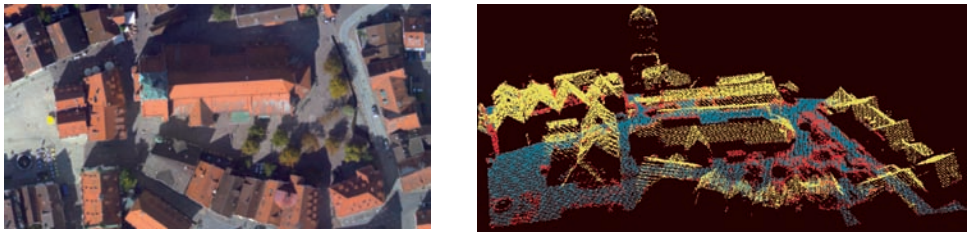

Fig. 6: Classification results in a dense urban area (Biberach). Left: orthoimage. Right: classified point cloud (same colours as in Fig. 5).

and artificial ground surfaces can be visually distinguished using such feature, the SVM classifier fails in splitting these classes. This result was partially expected since the histograms of α values for the two classes are very close (*cf.* Fig. 2). Another attribute has to be found to discriminate ground surfaces better.

Fig. 5 and 6 give examples of classified point over the two surveyed areas. Moreover, by merging the two terrain classes, the overall accuracy of the remaining three classes reaches 0.92 for the Biberach area. It shows that the SVM method is suitable for lidar point classification in dense build-up areas.

Similar accuracies have been reported for instance in (Matikainen et al. 2003, Vosselman et al. 2004), with multi-echo lidar data, sometimes used jointly with aerial images. It is not straightforward to compare the results since different data sets are classified using different methodologies. For this purpose, an optimal SVM classification using only "multi-echo" features has first to be performed on the data sets. Then, waveform features can be introduced to assess the real contribution of full-waveform lidar data. Moreover, since results strongly rely on the order of feature introduction, a feature discriminant analysis has also to be carried out to evaluate this contribution.

7 Conclusions and Perspectives

A flexible methodology for full-waveform lidar data analysis and classification in urban areas has been proposed in this article. In a first part, it has been shown that accurately modelling waveforms improves signal fitting and provides point clouds with additional useful parameters. Such parameters can physically be interpreted and significantly contribute to an appropriate classification algorithm. The main limitation is that the parametric expression of the waveform functions has no longer a simple formulation and new algorithms are needed to perform the optimization step. The Reversible Jump Markov Chain Monte-Carlo (RJMCMC) technique is one of them and will be soon used to handle more complex modelling functions.

In a second part, we can conclude that the SVM is a suitable methodology to perform classification in urban areas since it can handle classical geometric features like the 3D coordinates together with new features extracted from the waveform processing step. First results are promising; discrimination of buildings, vegetation, and ground regions was achieved with 92% accuracy in dense urban areas. Segmentation of different kind of surfaces is also possible.

Many improvements are conceivable with regards to the results. First, other generic SVM kernels have to be tested. On the other hand, a specific kernel can be formulated dedicated to our specific task. For that purpose, the number and kind of features has to be adapted and therefore synthetic cues must be found. Another solution is to iteratively process SVM classification focusing at each step on a specific class and segment it more precisely. A third possibility is to skip the step of feature choice and to use the vectors of the captured waveforms instead.

Finally, the classification results shall be the foundation of higher-level reasoning aiming at the 3D reconstruction of buildings. For this purpose geometric and topologic object features will be modelled, which are required for instance for object grouping.

Acknowledgments

The authors would like to thank the Maison de la Télédétection (French Remote Sensing Center) for providing the data set over Le Brusquet.

References

Bishop, C., 2006: Pattern Recognition and Machine Learning. – Springer, New York, USA.

Chapelle, O., Vapnik, V., Bousquet, O. & Mukherjee, S., 2002: Choosing Multiple Parameters for Support Vector Machines. – Machine Learning **46** (1): 131–159.

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M. & Puech, W., 2007: Processing full-waveform lidar data: modelling raw signal. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3/W52): 102–107.

Gross, H., Jutzi, B. & Thoennessen, U., 2007: Segmentation of tree regions using data from a full-waveform laser. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3/W49A): 57–62.

Haala, N. & Brenner, C., 1999: Extraction of buildings and trees in urban environments. – ISPRS Journal of Photogrammetry & Remote Sensing **54** (2/3): 130–137.

Harding, D., Lefsky M. & Parker, G., 2001: Laser altimeter canopy height profiles. Methods and validation for closed-canopy, broadleaf forests. – Remote Sensing of Environment **76** (9): 283–297.

Hsu, C.-W. & Lin, C.-J., 2002: A comparison of methods for multi-class Support Vector Machines. – IEEE Transactions on Neural Networks **13** (6): 415–425.

Jutzi, B. & Stilla, U., 2005: Waveform processing of laser pulses for reconstruction of surfaces in urban areas. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (8/W27).

Kirchhof, M., Jutzi, B. & Stilla, U., 2008: Iterative processing of laser scanning data by full waveform analysis. – ISPRS Journal of Photogrammetry & Remote Sensing **63** (1): 99–114.

Matikainen, L., Hyyppä, J. & Hyyppä, H., 2003: Automatic Detection of Buildings from Laser Scanner Data for Map Updating. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **34** (3/W13): 218–224.

Rottensteiner, F., Trinder, J., Clode, S. & Kubik, K., 2005: Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. – Information Fusion **6**: 283–300.

Schölkopf, B., Burges, C.J. C. & Smola, A.J., 1998: Advances in Kernel Methods. Support Vector Learning. – The M.I. T. Press, Cambridge, USA.

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 & Remote Sensing **59** (1–2): 85–101.

Tóvári, D. & Vögtle, T., 2004: Classification methods for 3D objects in laserscanning data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B3): 408–413.

Vosselman, G., Gorte, B.G. H. & Sithole, G., 2004: Change detection for updating medium scale maps using laser altimetry. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B3): 207–212.

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 & Remote Sensing **60** (2): 100–112.

Wagner, W., Ullrich, A., Melzer, T., Briese, C. & Kraus, K., 2004: From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B3): 201–206.

Xu, G. & Zhang, Z., 1996: Epipolar Geometry in stereo, motion and object recognition. – Kluwer Academic Publishers, Boston, USA.

Addresses of the Authors:

Clément Mallet and Dr. Fréderic Bretar, Laboratoire MATIS, Institut Géographique National, 2-4 av. Pasteur, F-94165 Saint-Mandé Cedex, Tel.: +33-1-43-98-80-00 + 75 66, Fax: +33-1-43-98-85-81, e-mail: clement.mallet|frederic.bretar@ign.fr

Prof. Dr.-Ing. Uwe Soergel, Institut für Photogrammetrie und GeoInformation (IPI), Leibniz Universität Hannover, Nienburger Straße 1, D-30167 Hannover, Tel.: +49-511/762-2981, Fax: +49-511/762-2483, e-mail: soergel@ipi.uni-hannover.de

Manuskript eingereicht: April 2008
Angenommen: Juni 2008