ANALYSIS OF FULL-WAVEFORM LIDAR DATA FOR CLASSIFICATION OF URBAN AREAS

Clément Mallet[1], **Uwe Soergel**[2], **Frédéric Bretar**[1]

[1] Laboratoire MATIS - Institut Géographique National
2-4 av. Pasteur, 94165 Saint-Mandé, FRANCE, firstname.lastname@ign.fr
[2] Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover
Nienburger Str.1 30167 Hannover GERMANY, soergel@ipi.uni-hannover.de

Commission III, WG III/3

KEY WORDS: lidar, signal processing, waveform analysis, modelling, classification, urban

ABSTRACT:

In contrast to conventional airborne multi-echo laser scanner systems, full-waveform (FW) lidar systems are able to record the entire emitted and backscattered signal of each laser pulse. Instead of clouds of individual 3D points, FW 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 FW data in urban areas. The problem of modelling FW lidar signals is first 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, the generalized Gaussian model is selected instead here. Then, a pattern recognition method for urban area classification is proposed. A supervised method using Support Vector Machines is performed on the FW 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.

1 INTRODUCTION

In the last decade, airborne lidar systems have become an alternative source for acquisition of altimetrer data. Such devices deliver a reliable, fast and accurate representation of terrestrial landscapes through georeferenced and unstructured 3D point clouds (RMSE $<$ 0.1 m in altimetry). Range is determined directly according to the signal runtime measurement whereas photogrammetric techniques derive the 3D information indirectly based on the camera orientations and the disparity of correspondences in stereo photos 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 (Matikainen et al., 2003; Sithole and Vosselman, 2004). They often depend on the availability of a cadastral map, even if, without this information, building outlines can at least roughly be extracted. 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 to be sensitive to large off-terrain objects and surface discontinuities. Therefore, many authors proposed other inputs like echo intensity (Tóvári and Vögtle, 2004) or multi-spectral images (Rottensteiner et al., 2005) to achieve better results.

Since few years, a new generation of lidar devices designed to digitize and to record the entire backscattered signal of each emitted laser pulse became operational. They are called **full-waveform (FW)** 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 FW data processing, mainly in vegetated areas. The higher point density inside the penetrated canopy offers insight in the vegetation types and state (Harding et al., 2001). In urban areas, the potential of such data has been barely investigated. For instance, Jutzi and Stilla (2005) extract linear features on roof tops by establishing neighbourhood relationships between waveforms. They also aim at localizing more accurate building outlines. On the other hand, by exploiting other features in addition to the geometry (*e.g.*, pulse amplitude or width), segmentation 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. The general outline of this work is described in Section 2. Then, a new modelling function is proposed to process raw signals in Section 3. The results of the integration of the previous extracted features into a supervised classification algorithm are presented in Section 4. The aim is to discriminate four classes: buildings, vegetation, artificial 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 finally 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, FW 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 FW lidar data to extract 3D point clouds featuring more useful parameters in addition to the traditional (x, y, z) coordinates and to perform subsequently a point classification based on these parameters.

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. The aim of waveform processing is to maximize the detection rate of relevant peaks in order to foster information extraction from the raw signal. Non-parametric methods like splines, neural networks or Parzen windows are known to work very well to fit 1D signals. Nevertheless, they eventually approximate curves to the data (Bishop, 2006) and do not provide the signal maxima location, which is required to produce 3D point clouds through a georeferencing process.
The objective of waveform processing is twofold. A parametric decomposition increases the accuracy of the signal maxima along the lidar beam. 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 is to assess whether or not each new feature introduced is relevant for classification and how significant it is for urban analysis (does it provide really useful information?). The features are fed into a supervised classification algorithm using Support Vector Machines (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 (Huang et al., 2002). SVM allows to use jointly classical geometric features (number of extracted peaks, altitude difference between first and last echoes of a waveform, etc.), image-based information (Secord and Zakhor, 2007) as well as in our case new parameters extracted from the post-processing step.
The methodology for classification in urban areas by FW lidar data 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 adjust the level of detail of the classification.

3 WAVEFORM PROCESSING

3.1 Methodology

Our methodology is based on a paper written by Chauve et al. (2007). The authors describe an iterative waveform processing using a Non-Linear Least Squares fitting algorithm. After coarse initial peak detection, missing peaks are found in the residuals of the difference between the modelled and initial signals. 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 graces 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 algorithm using thresholds.
Moreover, the methodology has been improved to take the 'ringing effect' into account: after the sampled emitted pulse, a small secondary maximum due to the effects of the hardware waveform processing chain can be seen. Consequently, in urban areas, when the laser beam hits plane objects of high reflectance and with a small angle of incidence (typically streets and roofs), such artifact is still present in the reflected waveform. It is typical of FW sensors and does not exist in multiple-pulse point clouds. In the iterative process, a weak pulse just behind a strong one is now removed when their amplitude ratio is closed to the ratio computed from the emitted pulse (given with the data).

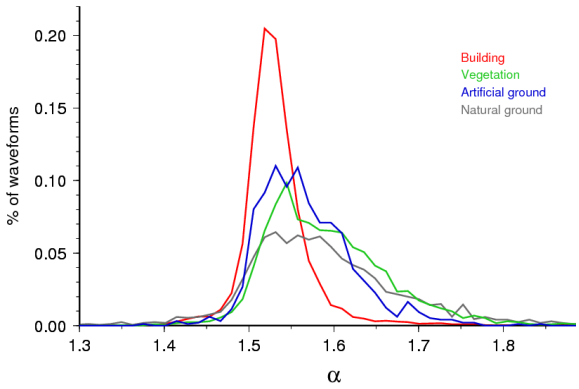

Figure 1: Histogram of α values over the four test classes.

3.2 Modelling functions

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 (Hofton et al., 2000). Wagner et al. (2006) have shown that more than 98% of the observed waveforms collected from the RIEGL system could be fitted with 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, in fact 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 figure 1. Hence, the characteristics of return peaks may differ significantly. It was already shown that standard extensions of the Gaussians model, which are 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). Here, the GG model was used also to process two FW data sets of different sites:

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

where A is the pulse amplitude, σ its width, μ the function mode and α a shape parameter which allows to simulate Gaussian ($\alpha = \sqrt{2}$), flattened ($\alpha > \sqrt{2}$) or peaked ($1 \leq \alpha < \sqrt{2}$) pulses.
A, σ and α are the three new features which will be introduced in the classification step in Section 4.
The Lognormal model fits asymmetric pulses with success but fails for symmetric ones.

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 character-

ize urban areas: buildings, vegetation, natural ground and artificial ground. *Artificial ground* gathers all kinds of streets (tar, asphalt,...) and pavements 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 $\mathcal{D}$ is the data space, $\mathcal{Y}$ the label space and A the training set (*e.g.*, $\mathcal{D} = \mathbb{R}^n$, $\mathcal{Y} = \{-1, 1\}$ in two-class problem). The classification is carried out using a linear discriminant function ω ($\mathcal{D} \rightarrow \mathcal{Y}$). $x_i \in \mathcal{A}$ 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, $\omega(x) = \mathbf{w} \cdot x - \theta$, where $\mathbf{w} \in \mathcal{D}$ is the normal vector to the hyperplane and θ the bias. We aim at finding the classifier parameters ($\mathbf{w}$, θ) which verify:

$$\forall (x_i, y_i) \in A, \quad y_i \times (\mathbf{w} \cdot x_i - \theta) > 0 \tag{2}$$

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

$$\forall (x_i, y_i) \in A, \quad y_i \times (\mathbf{w} \cdot x_i - \theta) \geq 1 \tag{3}$$

The support vectors lie on two hyperplanes $\mathbf{w} \cdot x - \theta = \pm 1$ 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 $2/\|\mathbf{w}\|$) 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. 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. It becomes:

$$\forall (x_i, y_i) \in A,\ y_i \times (\mathbf{w} \cdot x_i - \theta) > 1 - \xi_i \,/\, \forall i \in [1, N],\ \xi_i \geq 0 \tag{4}$$

The final optimization problem is subsequently:

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

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

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 $\Phi : \mathcal{D} \rightarrow \mathcal{H}$, in which the new distribution of samples enables the fitting of a linear hyperplane. Kernels 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 x_i points are implicitly projected in $\mathcal{H}$ by defining a kernel $K : \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}$ with $K(x_i, x_j) = (\Phi(x_i) \mid \Phi(x_j))$. 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).

Multi-class SVMs SVMs are designed to solve binary problems. When having $n \geq 3$ 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 and Lin, 2002). For such pairwise classification, $\frac{n\,(n-1)}{2}$ 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 `http://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 parameters on the pulse shape (or even strong hints about characteristic behaviours on urban areas), the design of a kernel dedicated to our specific purpose given our cues is a very difficult task. Therefore, a generic kernel was selected, the Gaussian kernel defined as:

$$K(x_i, x_j) = \exp\left(\frac{\|x_i - x_j\|^2}{2\gamma^2}\right) \text{ with } \gamma > 0 \tag{6}$$

where γ tunes 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 are not known beforehand, a grid search is performed in which the cross-validation accuracy (CVA) is computed for each point. In a v-fold cross-validation procedure, 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. The CVA represents the percentage of samples correctly classified averaged over all the subsets when they were used as the testing subset. The (C,γ) grid is composed of exponentially growing values of C and γ, for instance, in our study $C, \gamma = 2^{-15}, 2^{-13}, \ldots, 2^{15}$. After the coarse grid search, a finer one is computed in a smaller range around the optimal parameters found in the first step. Such grid search is necessary since the CVA over (C,γ) set is not convex.

4.4 Feature selection and relevancy

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

- Δr: difference between the pulse range and the highest range (lowest altitude) found in a large spherical environment (20m radius for instance),
- $\mathcal{R}$: residuals computed from a plane estimated by a robust L-estimators with norm $L_{1.2}$ (p=1.2 is proved to be the optimal value for the L_p estimator, see (Xu and Zhang, 1996) for more details) on the points in a given neighbourhood (here a spherical environment of 0.5 m radius),
- n_z: the deviation of the local normal vector from the vertical,
- Δz_{fl}: the altitude difference between the first and the last pulse of the waveform,
- $\mathcal{N}$: the number of echoes in the waveform,

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

- A, σ, α: the pulse amplitude, width, and shape respectively (extracted from the waveform processing step described in Section 3).

The three first parameters can be used with every 3D point cloud (only geometric information). The three last ones are derived by waveform modelling (amplitude can also be available with multiple-pulse point clouds).
Feature Δr allows to filter points on the terrain from off-ground points; Δz_{fl} and $\mathcal{N}$ discriminate vegetation points from the others. These two information are necessary because the number of echoes alone is not sufficient. Multiple reflections can occur when the laser beam hits a roof (due to superstructures) and the street (due to cars or building edges). $\mathcal{R}$ and n_z values are also affected by such data. The generalized Gaussian parameters are introduced in the SVMs to see how significant they are for the segmentation between the four classes and especially natural and artificial grounds.
Table 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.

↓ **Feature**	Building	Veget.	Art. grd	Nat. grd
Δr	variable	variable	$\rightarrow 0$	$\rightarrow 0$
$\mathcal{R}$	$\rightarrow 0$	high	$\rightarrow 0$	$\rightarrow 0$
n_z	[-45,45^o]	variable	[-10,10^o]	[-10,10^o]
Δz_{fl}	0	high	0	0
$\mathcal{N}$	$\sim$1	≥ 1	1	1
A	variable	medium	low	variable
σ	medium	high	variable	variable
α	[1.5,1.6]	variable	$\simeq \sqrt{2}$	$> \sqrt{2}$

Table 1: Empirical values of the selected features for SVM classification for the four labels (*Veget.*: vegetation, *Art. grd*: artificial ground, *Nat. grd*: natural ground).

5 FULL-WAVEFORM LIDAR DATA

Two data sets are available for this study. The data acquisitions were performed respectively in September 2006 and May 2007 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 each survey are described in Table 2. RIEGL full-waveform system allows to determine the vertical distribution of targets within the diffraction cone with a temporal sampling of 1 ns.
Each return waveform is composed of one or two sequences of 60 and 80 samples (for Biberach and Le Brusquet, respectively). 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 around 30m 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 an alpine rural region.

Area	Biberach	Le Brusquet
Urban specificity	dense	rural
Flight height (m)	500	700
Footprint size (m)	0.25	0.35
PRF (kHz)	100	111
Pulse width (ns)	≥ 5	
Temporal sampling (ns)	1	
Vertical section (m)	18 or 36	24 or 48
Pulse density (/m^2)	2.5	5

Table 2: Overview of the specification of the data sets.

6 RESULTS AND DISCUSSION

6.1 Modelling raw signals

As described in details 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 threefold.
First, the GG model improves signal fitting. More relevant points are extracted. 5% additional pulses are found which correspond to weak pulses in 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 Figure 2). On ground and building regions, ringing points are removed ($>$ 15% of the total number of points).
Furthermore, decomposing parametrically the waveforms increases the accuracy of the signal maxima location along the lidar beam. The target range detection is subsequently improved by more than 0.05 m on building roofs and ground.
Finally and above all, the global fitting quality is increased, because flattened, narrow, and high pulses are now well detected.

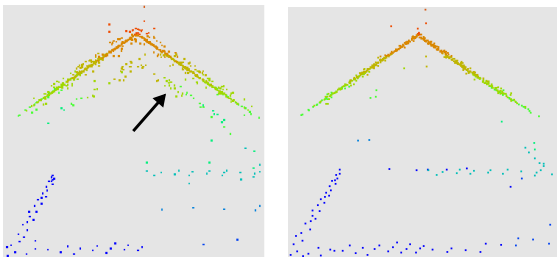

Figure 2: Building point cloud without taking the 'ringing effect' into account (**left**, the black arrow shows the false point layer.). The same data but after removal of artifacts (**right**). The roof appears no longer doubled.

Figure 1 shows that since the 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 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 a genuine information about the target shape and reflectance.

6.2 Behaviour of extracted parameters

A morphological analysis 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 the 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 an increasing roof slope tends to increase pulse width.
- Very low and high **shape** values are characteristics of building edges and vegetation. Building region corresponds with 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 Figure 1).

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. For all the tests carried out, the correct classification rate for the training step oscillates between 80 and 90%. It illustrates that the SVM classifier does not over-fit, but is able to generalize and has been trained sufficiently.
Table 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 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 (see table 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 conceivable when enough training samples are available (Figure 4). 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.

Area (number of reference points)	Buildings	Vegetation	Art. ground	Nat. ground
Building (76593)	**87.1**	8.8	3.6	0.5
Vegetation (8943)	10.2	**88.9**	0.7	0.2
Art. ground (49048)	2.2	2.1	**84.6**	11.1
Nat. ground (1043)	4.1	~ 0	33.2	**62.7**

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

The Overall Accuracy is used as a quality criterion and is defined as:

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

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/dim $\mathcal{Y}$ when the classifier randomly chooses the class for each point with the same probability. Figure 3 shows the evolution of the classification accuracy depending on the input features, adding them by their historical 'order of appearance'

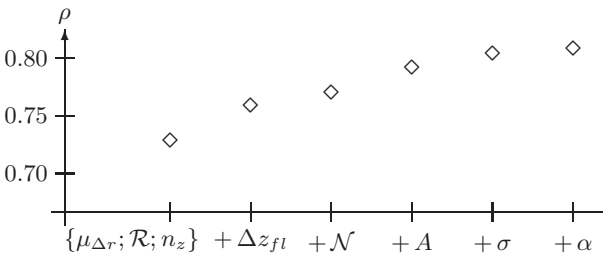

Figure 3: Overall Accuracy evolution depending on the features included in the SVM algorithm. Starting from the vector $\{\mu_{\Delta r}; \mathcal{R}; n_z\}$, the other ones are added progressively (Biberach area).

(see part 4.4). Each new feature improves the classification results. A label-by-label analysis reveals that the amplitude value allows to discriminate building and ground points; the feature σ is helpful to enhance the building/vegetation separation. In reality, 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 figure 4, no ground truth available for this area). Another solution has to be found to discriminate ground surfaces better.
The figures 4 and 5 give examples of classified point over the two

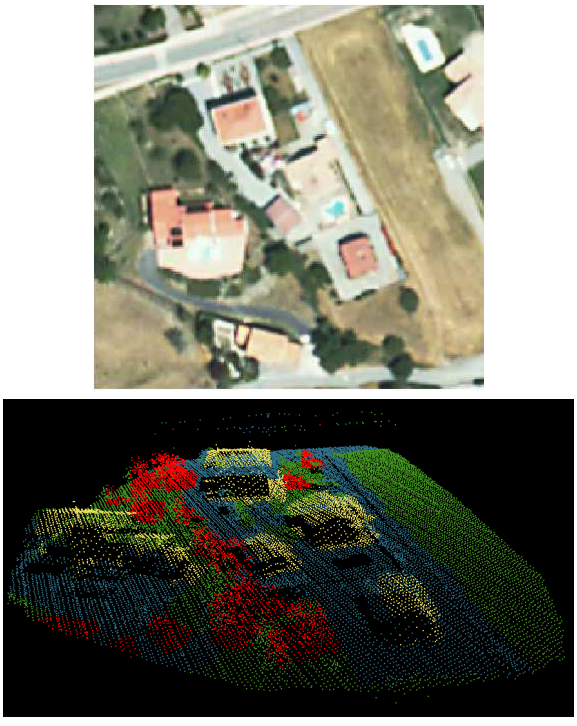

Figure 4: Classification results in a scattered urban area (Le Brusquet area). **Above**: orthoimage of the region of interest. **Below**: classified point cloud (yellow: buildings, red: vegetation, blue: artificial ground and green: natural ground).

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.

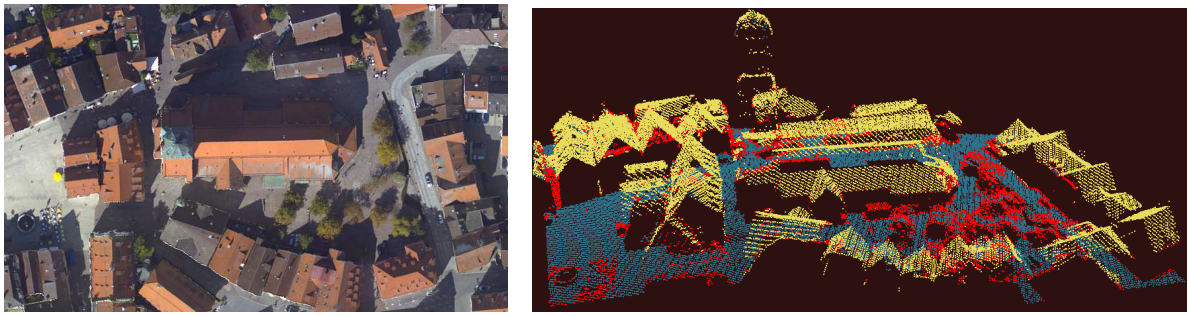

Figure 5: Classification results in a dense urban area (Biberach city). **Left**: orthoimage of the region of interest. **Right**: classified point cloud (yellow: buildings, red: vegetation, blue: artificial ground and green: natural ground).

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 modelling accurately waveforms improves signal fitting and provides point clouds with additional useful parameters. Such parameters are physically interpretable and significantly contribute to an appropriate classification algorithm. The main limitation is that the parametric expression of the waveform functions has no longer 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.

Similar accuracies have been reported for instance in (Matikainen et al., 2003), with only multi-echo lidar data. Classification with features used in such paper and FW features has to be performed to assess the real contribution of full-waveform lidar data.

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 our specific task. For that purpose, the number of features has to be reduced and therefore synthetic cues found. Another solution is perhaps to iteratively process SVM classification focusing at each step on a specific class and segment it more precisely. A third possibility is eventually to skip the step of feature choice and to use the vectors of the FW data 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.

References

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

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M. and Puech, W., 2007. Processing full-waveform lidar data: modelling raw signals. In: IAPRS, Vol. 39 (Part 3/W52), Espoo, Finland, pp. 102–107.

Gross, H., Jutzi, B. and Thoennessen, U., 2007. Segmentation of tree regions using data of a full-waveform laser. In: IAPRS, Vol. 36 (Part 3/W49A), Munich, Germany, pp. 57–62.

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

Hofton, M., Minster, J. and Blair, J., 2000. Decomposition of Laser Altimeter Waveforms. IEEE Trans. on Geoscience and Remote Sensing 38(4), pp. 1989–1996.

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

Huang, C., Davis, L. and Townshend, J., 2002. An assessment of support vector machines for land cover classification. International Journal of Remote Sensing 23(4), pp. 725–749.

Jutzi, B. and Stilla, U., 2005. Waveform processing of laser pulses for reconstruction of surfaces in urban areas. In: Urban 2005, Vol. 36 (Part 8/W27), Tempe, USA.

Matikainen, L., Hyyppä, H. and Hyyppä, J., 2003. Automatic Detection of Buildings from Laser Scanner Data for Map Updating. In: IAPRS, Vol. 34 (Part 3/W13), Dresden, Germany, pp. 218–224.

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, pp. 283–300.

Schölkopf, B., Burges, C. and Smola, A., 1998. Advances in Kernel Methods. Support Vector Learning. The MIT Press, Cambridge, USA.

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 & Remote Sensing 59(1-2), pp. 85–101.

Tóvári, D. and Vögtle, T., 2004. Classification methods for 3D objects in laserscanning data. In: IAPRS, Vol. 35 (Part B3), Istanbul, Turkey, pp. 408–413.

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.

Wagner, W., Ullrich, A., Melzer, T., Briese, C. and Kraus, K., 2004. From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges. In: IAPRS, Vol. 35 (Part B3), pp. 201–206.

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