CONTRIBUTION OF AIRBORNE FULL-WAVEFORM LIDAR AND IMAGE DATA FOR URBAN SCENE CLASSIFICATION

Nesrine Chehata[1,2], *Li Guo*[1]

1- Institut EGID, GHYMAC laboratory
1 allée F. Daguin 33670 Pessac, France.
E-mail: name.surname@egid.u-bordeaux3.fr

Clément Mallet[2]

2- IGN, MATIS laboratory
4 av Pasteur 94165 Saint-Mandé, France.
E-mail: name.surname@ign.fr

ABSTRACT

Airborne lidar systems have become an alternative source for the acquisition of altimeter data. In addition to multi-echo laser scanner systems, full-waveform systems are able to record the whole backscattered signal for each emitted laser pulse. These data provide more information about the structure and the physical properties of the surface. This paper is focused on the classification of full-waveform lidar and airborne image data on urban scenes. Random forests are used since they provide an accurate classification and run efficiently on large datasets. Moreover, they provide measures of variable importance for each class. This is crucial to analyze the relevance of each feature for the classification of urban scenes. Random Forests provide more accurate results than Support Vector Machines with an overall accuracy of 95.75%. The most relevant features show the contribution of lidar waveforms for classifying dense urban scenes and improve the classification accuracy for all classes.

1. INTRODUCTION

Airborne Laser Scanning (ALS) is an active remote sensing technique providing direct range measurements between the laser scanner and the earth topography. Range is determined directly from the signal runtime measurements. Such distance measurements are mapped into unstructured 3D point clouds using direct georeferencing (GPS/INS). The altimeter accuracy of a topographic lidar measurement is high (<0.1 m). Depending on the geometry of illuminated surfaces, several backscattered echoes can be recorded for a single pulse emission. Many authors showed the potential of multi-echo lidar data for urban area analysis and building extraction [1]. 3D point cloud classification can be based on geometric and textural attributes [2]. Other works include the lidar intensity [3] or combine lidar and multispectral data [4]. Since 2004, new commercial ALS systems called full-waveform (FW) lidar have emerged with the ability to record the complete waveform of the backscattered 1D-signal. Each echo in this signal corresponds to an encountered object. Thus, in addition to range measurements, further physical properties of objects included in the diffraction cone may be revealed by analyzing the shape of backscattered waveforms. A detailed state-of-the-art on full-waveform topographic lidar can be found in [5]. In urban scenes, the potential of such data has been barely investigated. In [6, 7], FW lidar features are used , in addition to the geometry, to detect vegetated areas.
In this work, different FW lidar and image features are extracted to classify urban scenes into four classes: buildings, vegetation, natural ground, and artificial ground. Artificial ground gathers all kinds of streets and street items such as cars, traffic lights whereas the natural ground includes grass, sand, and bare-earth regions. The objective is to provide an accurate classification using Random Forests and to measure the contribution of full-waveform lidar features for classification purposes on urban scenes.

This paper is organized as follows. The lidar and optical features will be introduced in Section 2. In Section 3, the classification process by Random Forests is presented. Experimental results are then given in Section 4, and finally, conclusions are drawn in Section 5.

2. FULL-WAVEFORM LIDAR AND AIRBORNE IMAGE DATA FEATURES

Our multisource data are composed of full-waveform lidar data and a color airborne orthoimage. In order to combine lidar and image data, lidar points are projected into 2D image geometry using a bilinear interpolation (*cf.* Figure 1). For each pixel, Red, Green and Blue channels of the orthoimage are used as three independent features. Besides, specific lidar features are detailed in Section 2.
Thus, our feature vector is composed of twelve components: **three** optical components R, G and B, **five** multi-echo lidar components and **four** FW lidar components. The resulting feature vector fv for each site is given by:

$$fv = [R\ G\ B; \Delta z\ N_z\ \mathcal{R}_z\ N\ N_e; A\ w\ \sigma\ \alpha]^T \quad (1)$$

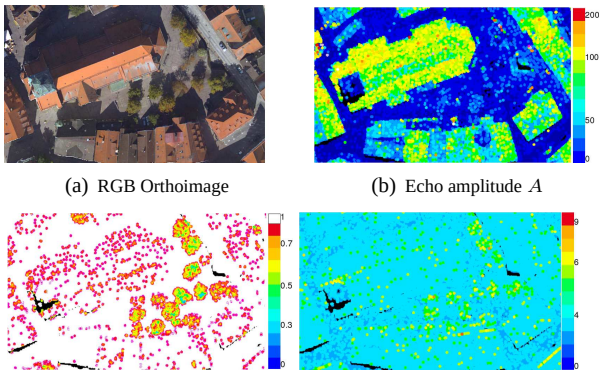
The five multi-echo lidar features we consider are spatial and can be processed from either multi-echo or full-waveform lidar data. These characteristics are computed using a volumetric approach within a local neighborhood ν_P at each lidar point P. The local neighborhood includes all the lidar points within a cylinder, with a fixed radius, centered at the point. In urban scenes, most objects can be described by planar surfaces such as building roofs and roads. The planarity of the local neighborhood will help discriminating buildings from vegetation. The local plane Π_P is estimated using a robust M-estimator with norm $L_{1.2}$ [8] on the points in ν_P. The features are:

- Δz: height difference between the lidar point and the lowest point found in a large cylindrical volume whose radius has been experimentally set to 15 m. This feature will help discriminating ground and off-ground objects.
- N_z: deviation angle of the local normal vector (related to the local plane Π_P) from the vertical direction. This feature highlights the ground.
- $\mathcal{R}_z$: residuals of the local plane estimated in a small cylinder (0.5 m radius). Residuals should be high for vegetation.
- N: total number of echoes within the waveform of the current lidar point. This feature will be high for vegetation and building facades.

- N_e: normalized number of echoes obtained by dividing the echo number by the total number of echoes within the waveform of the current lidar point. This feature highlights the vegetation since multiple reflections can occur on it (*cf.* Figure 1).

The remaining features are more specific to FW lidar data and are obtained by modelling the lidar waveforms. The amplitude and Echo width are described in [9].

- A: echo amplitude. High amplitude values can be found on building roofs, on gravel, and on cars. Asphalt and tar streets have a low values The lowest values correspond to vegetation due to a higher target heterogeneity and attenuation (*cf.* Figure 1).
- w: echo width. Higher values correspond to vegetation since it spreads the lidar pulses. A narrow pulse is likely to correspond to ground and buildings. However, the width value may increase with roof slope (*cf.* Figure 1).
- σ: echo cross-section, equals to $A \times w$. The values are high for buildings, medium for vegetation and low for artificial ground.
- α: echo shape describing how locally distorted the waveform is. [10] shows that very low and high shape values correspond relatively to building roofs and vegetation.



(a) RGB Orthoimage (b) Echo amplitude A

(c) Normalized number of echoes N_e (d) Echo width w

Fig. 1. Orthoimage and examples of FW lidar features.

Table 1 summarizes the expected lidar feature values for the different classes we consider on urban scenes.

Lidar	Feat.	Building	Vegetation	Art.Grnd	Nat. Grnd
ME	Δz	variable	variable	$\rightarrow 0$	$\rightarrow 0$
	N_z	[-45,45^o]	variable	[-10,10^o]	[-10,10^o]
	$\mathcal{R}_z$	$\rightarrow 0$	high	$\rightarrow 0$	$\rightarrow 0$
	N	1	≥ 1	$\sim$1	1
	N_e	1	≤ 1	$\sim$1	1
FW	A	variable	medium	low	variable
	w	medium	high	variable	variable
	σ	high	medium	low	variable
	α	[1.5,1.6]	variable	$\simeq \sqrt{2}$	$> \sqrt{2}$

Table 1. Empirical values of lidar features for the four classes.

3. RANDOM FORESTS

Random Forests is a variant of bagging proposed by Breiman [11]. It is a decision tree based ensemble classifier that can achieve a classification accuracy comparable to boosting [11], even Support Vector Machines (SVMs) [12]. It does not over fit, runs fast and efficiently on large data sets and can handle thousands of input variables without variable deletion. In addition, the significance of each variable in the classification can be estimated. These outstanding features make it suitable for the classification of remote sensing data such as multispectral data [12] or multisource data [13]. In this work, Random Forests are applied to classify airborne multisource data which are composed of airborne lidar and image data on urban scenes. Random Forests have not yet been used to such data.

Random Forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest [11]. As in Breiman's method, in training, the algorithm creates multiple bootstrapped samples of the original training data, then builds a number of no pruning Classification and Regression Trees (CART) from each bootstrapped samples set and only a randomly selected subset of the input variables is used to split each node of CART. For classification, each tree in the Random Forests gives a unit vote for the most popular class at each input instance. The label of input instance is determined by a majority vote of the trees. The number of variables M randomly chosen at each split is considered as the single user-defined adjustable parameter. This parameter is not critical and is often set to the square root of the number of inputs.

4. EXPERIMENTAL RESULTS

4.1. Data set

The data acquisition was carried out with the RIEGL LMS-Q560 system over the city of Biberach (Germany). The flight height was 500 m. The lidar point cloud has a point density of approximatively 2.5 pts/m^2 with a footprint size of 0.25 m. The orthophotography has been captured with an Applanix DSS 22M device. Its resolution is 0.25 m.

The city of Biberach includes artificial grounds, natural grounds, vegetation and buildings. The number of available reference samples is 797364 and they are split almost evenly between training and test samples. The ground truth is processed manually, based on a region oversegmentation of the orthophoto. This lidar data set has been used for classifying urban scenes using SVM [14]. Classification was processed for each 3D point separately. In this work, lidar features are considered in 2D geometry with optical image data.

4.2. Classification results

The Random Forests implementation software by L. Breiman and A. Cutler (`http://www.r-project.org`) was used in experiments. Figure 2 shows the classification result. Underlying parameters have been fixed to $M = 3$ which means that three variables are considered at each split and the number of trees was set to 60.

One can observe that errors occur mainly on building edges where lidar points correspond to a transition between building and artificial ground classes. Besides, these confusion errors are amplified due to the interpolation process of lidar points in 2D geometry. Other errors occur on artificial ground, essentially in shadowed streets: they are due to the use of RGB channels that have uniform irrelevant values in urban corridors.

The confusion matrix for test data is given in Table 2. The rows of the matrix are actual classes, and the columns are the predicted classes. The training dataset is highly imbalanced. We can notice that artificial ground and buildings are well classified with lower error rate. However, the algorithm has more difficulties in classifying natural ground and vegetation which suffer from smaller training set. As for vegetation, confusions essentially occur with artificial ground due to the lidar data interpolation. In fact, in non-dense vegetated areas, the lidar beam is likely to reach the ground underneath and the resulting waveform has mixed properties.

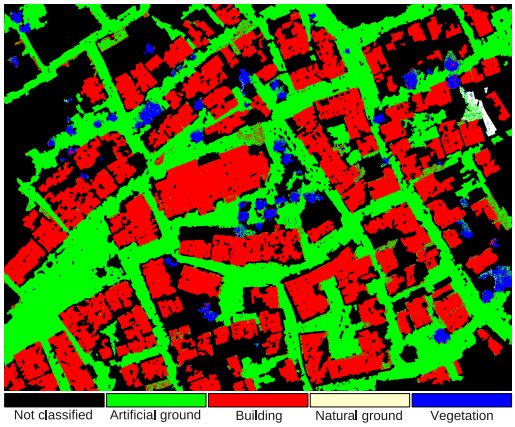


Fig. 2. Classification result (T=60 trees and M=3).

Class	Art.Grnd	Building	Nat.Grnd	Veget.	Error %
Art.Grnd	186952	5033	34	925	3.1
Building	6762	180903	8	342	3.8
Nat.Grnd	680	16	1409	44	34.4
Vegetation	2429	554	108	12632	19.6

Table 2. Confusion matrix for test data using 60 trees and 3 split variables. Total error rate=4.25%.

Random Forests results were compared to Support Vector Machines (SVM) ones. The algorithm was first proposed by [15] and has been used in much work in classification. Parameters were set to default. SVM were applied using two different kernels. $x_{i,j}$ are the data to be classified.

- polynomial: $K(x_i, x_j) = (\gamma\ {x_i}^T x_j)^3$.
- radial or Gaussian :$K(x_i, x_j) = e^{(-\gamma|x_i - x_j|^2)}$. $\gamma = 1/N$ where N is the data dimension.

Table 3 illustrates the confusion matrix for SVM classifier with a radial kernel. The error rates are higher than those of Random Forests. However, in both tables, the classifier behaves similarly with regard to different classes.

Class	Art.Grnd	Build.	Nat.Grnd	Veget.	Error %
Art.Grnd	182733	8822	36	1353	5.3
Building	9858	177732	5	420	5.5
Nat.Grnd	1140	14	954	41	55.6
Vegetation	3471	612	198	11442	27

Table 3. Confusion matrix for test data using SVM with a radial kernel. Total error rate=6.5%.

Table 4 compares Random Forests and SVM results.Random Forests gives better classification accuracy for all classes, in both cases with a total error rate of 4.25%. Moreover, training and test computing time are really improved. However a considerable amount of memory is needed to store a N by T matrix in memory.

4.3. Variable importance results

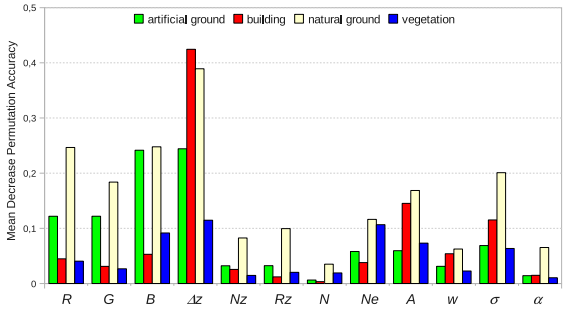
Aside from classification, Random Forests provide measures of variable importance for each variable.

4.3.1. Permutation importance measure

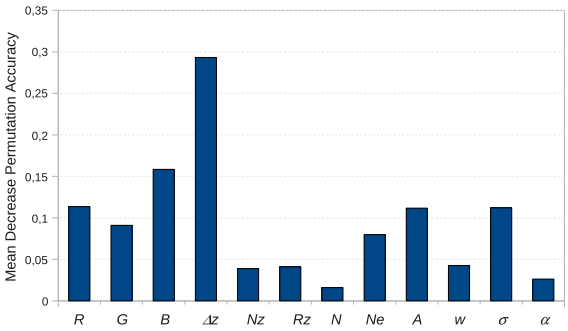
The importance measure is based on the permutation importance measure. When the training set for a particular tree is drawn by sampling with replacement, about one-third of the cases are left out of the sample set. These Out of Bag (OOB) data can be used to estimate the test accuracy and permutation importance measure. The importance of variable m can be estimated by randomly permuting all the values of the m^{th} variable in the OOB samples for each tree. A measure of variable importance can be a difference between prediction accuracy (*i.e.* the number of observations correctly classified) before and after permutating variable m, averaged over all trees [11]. A high prediction accuracy decrease indicates the importance of that variable. A variable importance estimate for our feature vector, using a balanced training data, is depicted on Figure 3 for each class.

Classifier	SVM polynomial	SVM radial	RF $T = 60\ M = 3$
Training time(s)	23107	15277	514
Test time(s)	3800	5347	14
Total Error rate	7.1%	6.5%	4.2%

Table 4. Comparison of Random Forests and SVM results.



(a) Variable importance per class.



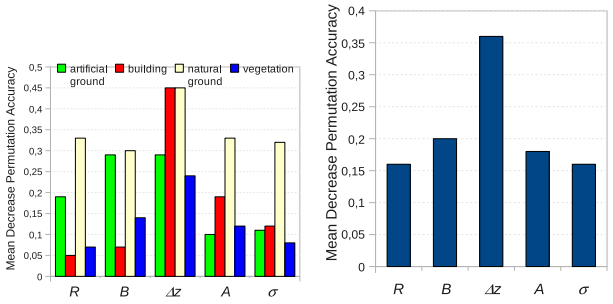
(b) Average variable importance.

Fig. 3. Variable importance by mean decrease permutation accuracy. Balanced training data with 3000 samples per class.

4.3.2. Discussion

From Figure 3, it is obvious that the height difference $\Delta\mathbf{z}$ is the most important feature for all classes. It is the only topographic feature. On the contrary, the number of echoes N is the less important feature as predicted since the number of echoes equals 1 for building roofs and grounds and is variable on vegetation. However, the normalized number of echoes is very discriminative for vegetation as showed in Table 1. For the building class, the most important variables are Δz and two other FW features: A and σ. RGB channels are ambiguous on roofs due to shadows. Both features that are related to

the planarity of the local neighboorhood (N_z, $\mathcal{R}_z$) are very sensitive to the slope. These features are not homogeneous for the building class. However, they should be relevant for roof segmentation. On the other hand, variable importance on both vegetation and natural ground classes are more dispersed between topographic, radiometric and some lidar features which makes them harder to classify (*cf.* Tables 2,3). Finally, the echo shape feature seems to be unuseful for the classification of urban scenes. Low values correspond to natural ground class but other lidar features such as Δz, σ or N_z are more relevant for this class. The most relevant features for all classes (*cf.* Figure 3(b)) are the height difference Δz, Red and blue channels R, B, the echo amplitude A, and the echo cross-section σ which confirms the contribution of optical images and FW lidar features to urban scene classification. The random forests were re-run with the most important features. Figure 4 depicts the variable importance that is highly similar to previously, which confirms the reliability of this measure. Moreover, the classification accuracy of all classes is enhanced using the five relevant features. The feature importance measure, using a balanced training data, is essential to select the best features for minor classes. Moreover, natural ground class is better classified than vegetation since the kept features are less important for the latter class.



(a) Variable importance per class. (b) Average variable importance.

Fig. 4. Variable importance for the most relevant features.

Class	Art.Grnd	Building	Nat.Grnd	Veget.	Error %
Art.Grnd	188137	3846	63	898	2.5
Building	4161	183493	6	355	2.4
Nat.Grnd	184	14	1943	8	9.6
Vegetation	1664	532	32	13495	14.2

Table 5. Confusion matrix for test data using best features, 60 trees, and 3 split variables. Total error rate=2.95%.

5. CONCLUSION

In this work, Random Forests were successfully applied to the classification and variable importance measure of a multisource remote sensing data set. In our experiments, the Random Forests classifier performed better than SVM, achieving a low error rate of 4.25% while enhancing training and test computing time. Classification results were very satisfactory for both building and artificial ground classes. Besides the algorithm is a user-friendly method with easy adjustable parameters. Furthermore, we successfully used this powerful method to estimate the importance of FW lidar and optical image features to classify urban scenes. The permutation accuracy criteria revealed that the most significant attribute is a topographic one: the relative height of a lidar point. followed by blue channel, two FW lidar features: the echo amplitude and the cross-section, and finally the red channel. Variable importance was carried out per class which enables to choose the best features for a specific application but also for all classes for a cartography purpose. Other lidar features such as $N_z, \mathcal{R}_z$ related to planarity should be more useful for roof segmentation. Provided more FW lidar attributes, the classification can be refined by dividing vegetation into tree species or segmenting building rooftops.
The final classification can be enhanced using spatial homogeneity or object-based segmentation in a regularization framework.

6. REFERENCES

[1] C. Frueh, S. Jain, and A. Zakhor, "Data Processing Algorithms for Generating Textured 3D Facade Meshes from Laser Scans and Camera Images," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 159–184, 2005.

[2] L. Matikainen, J. Hyyppä, and H. Hyyppä, "Automatic detection of buildings from laser scanner data for map updating," *International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences*, vol. XXXIII, no. 3/W13, pp. 218–224, 2003.

[3] A. Charaniya, R. Manduchi, and S. Lodha, "Supervised parametric classification of aerial lidar data," in *Real-Time 3D Sensors and their use workshop, in conjunction with IEEE CVPR*, 2004, p. 8p.

[4] J. Secord and A. Zakhor, "Tree detection in aerial lidar and image data," in *ICIP*, 2006, pp. 2317–2020.

[5] C. Mallet and F. Bretar, "Full-waveform topographic lidar: State-of-the-art," *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 64, no. 1, pp. 1–16, 2009.

[6] H. Gross, B. Jutzi, and U. Thoennessen, "Segmentation of tree regions using data of a full-waveform laser," in *Symposium of ISPRS Photogrammetric Image Analysis (PIA)*, Munchen, Germany, sep 2007, ISPRS, vol. XXXVI Part(3/W49A).

[7] W. Wagner, M. Hollaus, C. Briese, and V. Ducic, "3D vegetation mapping using small-footprint full-waveform airborne laser scanners," *International Journal of Remote Sensing*, vol. 29, no. 5, pp. 1433–1452, 2008.

[8] G. Xu and Z. Zhang, *Epipolar Geometry in stereo, motion and object recognition*, Kluwer Academic Publishers, 1996.

[9] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 60, no. 2, pp. 100–112, Jan. 2006.

[10] A. Chauve, C. Mallet, F. Bretar, S. Durrieu, M. Pierrot-Deseilligny, and W. Puech, "Processing full-waveform lidar data: modelling raw signals," in *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Espoo, Finland, 2007, vol. 36 (Part 3/W52), pp. 102–107.

[11] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[12] M. Pal, "Random Forest classifier for remote sensing classification," *International Journal of Remote Sensing*, vol. 26, no. 1, pp. 217–222, 2005.

[13] P.O. Gislason, J.A. Benediktsson, and J.R. Sveinsson, "Random forests for land cover classification," *Pattern Recognition Letters*, vol. 27, no. 4, pp. 294–300, 2006.

[14] C. Mallet, F. Bretar, and U. Soergel, "Analysis of full-waveform lidar data for classification of urban areas," *Photogrammetrie Fernerkundung GeoInformation (PFG)*, vol. 5, pp. 337–349, 2008.

[15] B.E. Boser, "A training algorithm for optimal margin classifiers," in *the 5th Annual ACM Workshop on Computational Learning Theory*, 1992.