Automatic LOD0 classification of airborne LiDAR data in urban and non-urban areas

J. Balado*, L. Díaz-Vilariño*, P. Arias and L. M. González-Desantos**

*Applied Geotechnologies Group, Department of Natural Resources and Environmental Engineering, School of Mining and Energy Engineering, University of Vigo, Vigo, Spain; **Department of Civil Engineering, University of Porto, Porto, Portugal

ABSTRACT

Point clouds are a very detailed and accurate vector data model of 3D geographic information. In contrast to other data models, no standard has been defined for visualization and management of point clouds based on levels of detail (LOD). This paper proposes the application of the concept of LODs to point clouds and defines the LOD0 for point cloud classification (the lowest possible level of detail) as urban and non-urban. A methodology based on the use of machine learning techniques is developed to perform LOD0 classification to airborne LiDAR data. Point clouds acquired with airborne laser scanner (ALS) are structured in grid maps and geometric features related with Z distribution and roughness are extracted from each cell. Six machine learning classifiers have been trained with datasets including urban (cities) and non-urban samples (farmlands and forests). The influence of grid size, point density, number of features and classifier type are analysed in detail. The classifiers have been tested in three case studies. The best results correspond to a grid size of 100 m and the use of 12 geometric features. The accuracy is around 90% in all tests and Cohen’s Kappa index reaches 81% in the best of cases.

Introduction

Point clouds focus great research in recent years. The technology responsible for their acquisition has become cheaper and more companies offer these services. In addition, the number of countries that offer LiDAR information in their cartographic databases along with classical photogrammetry is increasing (Centro Nacional de Información Geográfica, n.d.; Department For Environment Food & Rural Affairs, n.d.; Netherlands eScience Center, n.d.), since point clouds provide valuable 3D geometric information not available in the images. Even so, the structuring of LiDAR data is still a problem today despite the improved performance of computers and increased capacity in hard drives.

Although point clouds can be considered as vector data models of 3D points, they are not usually integrated into Geographic Information System (GIS) due to their large size and non-structured nature. The visualization of such magnitude of points becomes confusing to the human eye and presents an excessive data load for a computer. The difficulty of the work is closely related to the type of cloud with which to work. There are different types of point clouds depending on the equipment with which they were acquired, which entail different densities of information and different purposes. Point clouds acquired with a Terrestrial Laser Scanner (TLS) have higher density and serves to perform jobs that require great precision and where other kind of equipment does not reach (Lague, Brodu, & Leroux, 2013). Mobile Laser Scanner (MLS) has lower density than TLS but allows faster acquisition (million points per second), being limited to acquiring the environment near to the MLS trajectory, commonly near to roads (Kumar, Lewis, & McCarthy, 2017). Finally, Aerial Laser Scanner (ALS) has less density than the previous ones, but allows to acquire large areas of land quickly and from top view perspective (Yan, Shaker, & El-Ashmawy, 2015). Together, all these technologies can provide the user a complete point cloud of a scene with different densities.

There are standards of representation of geospatial information, either 2D (standard mapping in GIS) or 3D (CityGML (Biljecki, Ledoux, & Stoter, 2016)), which structure the information according to different levels of detail (LOD), allowing an optimal management and visualization information at the desired scale. For example, five LODs are defined in CityGML (LOD 0 to 4), being LOD0 the coarsest level, representing a building as a simple 2D polygon, and LOD4 the most complex, representing the building in a 3D polygon with all its external elements and even its internal furniture. However, although point clouds are vector models as others, there is no standardization and classification of
them according to LODs, essential to manage large amounts of data.

In order to be able to structure the data from a lower LOD to a higher one, this paper presents a LOD0 classification methodology for point clouds on a large-scale based on machine learning techniques. In point clouds, LOD0 has been defined as a classification in urban and not urban, although there are already methodologies that focus on the classification and detection of elements within urban areas on a street or ground scale with a fine LOD: (object classification (Golovinskiy, Kim, & Funkhouser, 2009; Weinmann, Jutzi, Hinz, & Mallet, 2015), classification of ground elements (Balado, Díaz-Vilarino, Arias, & González-Jorge, 2018), curb detection (Serna & Marcotegui, 2013; Vosselman & Liang, 2009), façade detection (Serna, Marcotegui, & Hernández, 2016). The classification presented in this paper is intended for large land areas with a coarse LOD, equivalent to a LOD0, necessary for the input data selection of the previous examples.

The definition of urban areas is ambiguous (Census Bureau United States, 2015; Schmidheiny & Susedekum, 2015), and in many cases, it is related to the number of inhabitants, population density or administrative criteria or political boundaries (United Nations Statistics Division, 2006), parameters not measurable in a point cloud. By contrast, urban areas have in common a high concentration of infrastructures and a distinctive geometry, which are measurable with the laser scanner.

To achieve the objective of the paper, ALS point clouds are used, which are the ones that have a lower density within the different acquisition methods and those that cover extensive areas. The methodology is based on a grid map to collect the data and extract the features related to Z distribution and roughness. Then, the features are used to train six machine learning classifiers: Trees, Logistic Regression, SVM, SVM with Gaussian distribution, KNN and Ensemble Bagged Trees. In the results section, the attributes (grid size, point density and number of features) are evaluated and the optimum classifiers are tested in three real case studies.

This paper is organized as follows. Related work about land classification, both based on images and LIDAR data, with urban areas are collected in Section 2. In Section 3, the datasets used are detailed. The methodology is explained in Section 4. In Section 5, the results of the methodology are analysed. Finally, Section 6 concludes this work.

Related work

The defined LOD0 proposed in this work for point cloud classification in urban and non-urban areas is related with Land Use and Land Cover (LULC) classification, where urban areas are defined (Anderson, Hardy, Roach, Witmer, & Peck, 1976), as well as other levels of classification (Levels I, II and III) of urban areas (residential, commercial, industrial and transportation) and non-urban areas (agricultural, rangeland, forest land, water, wetland, barren land, tundra and snow-ice); and its internal divisions. Most of works about urban and non-urban classification are defined and are focused on satellite images, while very few works use LIDAR data and even less 3D geometry in large amount data.

Urban classification and detection on satellite images

Urban land classification and detection is a well-known topic by now. There have been extensive studies of feature extraction in images for land classification. In most of them, methodologies have been tested in their own case studies, while some works have established a comparison with other methods by considering the same study area and images of the same satellite (Li, Wang, Wang, Hu, & Gong, 2014; Qian, Zhou, Yan, Li, & Han, 2015). Variations in the environment between seasons and weather conditions also influence the results (Saadat et al., 2011). Multi-temporal images are normally used to detect changes, mainly as deforestation and urban growth (Piedelobo et al., 2018). Detection of changes implies to know the previous land uses and to detect their later variations.

Classification studies based on satellite images are typically carried out for higher levels such as Levels I and II. Huang et al. (2017) classify the extensive area in Beijing (China) in three land types (vegetation, built-up and water) through Normalized Difference Vegetation Index (NDVI) trajectory using all Landsat images in Google Earth Engine. Jebur, Mohd Shafri, Pradhan, and Tehrany (2014) they focus their study within the urban areas, comparing pixel and object-based land cover classifications, obtaining a better result with the latter. Stow, Lopez, Lippitt, Hinton, and Weeks (2007) distinguish low socioeconomic residential areas from non-residential. Poursanidis, Chrysoulakis, and Mitraka (2015) make a comparison between Landsat 5 and 8 based on a land cover mapping of urban and peri-urban areas. Fu and Weng (2016) relate urbanization through land use and land cover changes to the temperature obtained by the Landsat imagery. Momeni, Aplin, and Boyd (2016) map complex urban areas and analyze how factors such as resolution and bands influence the classification. Lu and Weng (2006) address the classification of commercial, industrial, transportation, residential and non-urbans lands on impervious surface and population density. Castelluccio, Poggi, Sansone, and Verdoliva (2015) use neural networks
algorithms for land use classification, including the residential class among others.

**Lidar data in land classification**

The use of LIDAR data in land classification is not as widespread and studied as the use of satellite images, mainly because LIDAR data are not available for as long as satellite images. The main advantage of the use of LIDAR data lies in the direct availability of 3D geometry, which is not possible when using satellite images. Furthermore, the intensity/reflectivity of materials and the number of returns produced by the laser are very useful for characterizing the elements on the surface of the Earth.

The joint use of LIDAR data and satellite images for LULC classification can be done in two ways: simultaneous and hierarchical (Zhang & Lin, 2017). The simultaneous use combines extracted features from LIDAR, height and intensity, with multispectral images at the same time, associating LIDAR points to pixels (Cao, Wei, Zhao, & Li, 2012; Hartfield, Landau, & van Leeuwen, 2011). It results in a greater number of useful features for classification. Salah, Trinder, and Shaker (2009) extract 22 features by combining LIDAR and images. Also it is possible the generation of new features, for example, digital surface models (DSM) (Moreira, Fernandez, Pereira, & Moreira, 2017; Zhang, Qin, Huang, Fang, & Liu, 2015). Even data from other sensors such as Synthetic Aperture Radar can be added (Gamba & Houshmand, 2002).

Features extracted by the simultaneous use favour the use of machine learning techniques, for example Support Vector Machines (Zarea & Mohammadhadeh, 2016) or Random Forests (Guo, Chehata, Mallet, & Boukir, 2011) and also in unsupervised classification (Gerke & Xiao, 2014). Many studies show that the use of LIDAR data improves the accuracy of the results in comparison with the use of only images, although the percentage of improvement depends on the method: 9.1% (Luo et al., 2016), 18% versus only multispectral and 28% versus RGB images (Huang, Shuye, Lee, & Kao, 2008) and 38% (Salah et al., 2009).

In the hierarchical classification, the LiDAR data are used in different phases for classification. Typically, LIDAR data are used to segment the images in a first stage and then applied in a classification based on images. Yu, Liu, Zhang, and Wu (2009) uses satellite images to classify them in: water, vegetation and impervious surfaces, in a first stage. In a second stage, they use LIDAR data to extract height (distance to ground) and roughness features (standard deviation of heights) with the aim of classifying vegetation in different types. Creating a normalised DSM, they differentiate impervious surfaces: ground, ordinary buildings, high-rise buildings, and skyscrapers. The same principle can be used in a more complex way, alternating and combining LIDAR data and images in different steps of a classification (Guan, Ji, Zhong, Li, & Ren, 2013).

There are few works that only use LIDAR data for classification. Brennan and Webster (2006) use height and intensity to land cover classification resulting in 10 classes of a coastal environment with urban, mixed forest, and wetland-estuary. In an urban context, the intensity of airborne LIDAR data is employed to differentiate between trees, buildings, ground and low vegetation (Hui, Di, Xianfeng, & Deren, 2008). Other common use is building extraction. Priestnall, Jaafar, and Duncan (2000) develop an unsupervised classification to difference between bare land, trees and buildings; as part of a methodology to extract DEM (Digital Elevation Model) from LIDAR DSM. Rottensteiner and Briese (2002) use a hierarchic application of robust interpolation to extract a DTM that, compared with a DSM, allows to extract the buildings. DSMs created from airborne LiDAR data can be used in different temporal instants for change detection in urban growth (Teo & Shih, 2013) and more in detail to detect changes in buildings (Pang, Hu, Wang, & Lu, 2014).

In recent years, multispectral LIDAR technology has made it possible to combine the benefits of LIDAR data and multispectral imagery. Most works on the subject use the Titan Optech (Scaioni et al., 2018), which acquires three point clouds with different wavelengths: 1550 nm (IR – Channel 1), 1064 nm (NIR – Channel 2), and 532 nm (Green – Channel 3). The combination of the three clouds allows the extraction of different indexes, the most common being the pseudo NDVI (pseudo Normalized Difference Vegetation Index). They are very useful for the differentiation of types of vegetation and constructed elements (Wichmann et al., 2015), and it allows to implement a 3D land cover classification (Morsy, Shaker, El-Rabbany, & LaRocque, 2016; Zou, Zhao, Li, Yang, & Fang, 2016). The information extracted from the Lidar multispectral data can also be used for machine learning classifiers (Teo & Wu, 2017).

With regard to previous approaches, the authors present a methodology to automatically classify point clouds in urban and non-urban areas, defined in a LOD0 classification. The classification only uses geometric features (height and roughness) from airborne LIDAR point clouds, and it is tested in three case studies. To the best of our knowledge, no papers have been found focusing only on geometric features for classification at this LOD, as the geometric features do not seem to be of relevance for this level.
Data and land types

All points clouds used in this work are public and obtained from the download site of the “Instituto Geográfico Nacional” (Centro Nacional de Información Geográfica, n.d.; Instituto Geográfico Nacional, n.d.). The density of the point clouds is 0.5 points/m² and they are enriched with RGB and intensity information, only used for the visualization of figures.

As mentioned before, two land types are considered in this work: urban and non-urban. Urban zones are those with high density of buildings and infrastructures, such as cities, towns and industrial centres; and non-urban zones are those with low or without population density neither infrastructures, mainly forests, farmlands or zones with a small number of buildings and infrastructures. The data employed to train the machine learning classifiers consist of seven datasets (Table 1 and Figure 1). The data employed to test the methodology consist of three datasets (Table 2 and Figure 2). They are selected for the clarity of discern between urban and non-urban to the ground truth data. The case studies are part of Jerez de La Frontera and Zaragoza cities (case studies 1 and 2) and the town of Medina de Rioseco (case study 3) with their surroundings. All datasets, both training and testing, are located in Spain. Points of each case study are manually labelled in urban and non-urban classes according to density of buildings and infrastructures mentioned in the previous paragraph.

Table 1. Summary of datasets used to train the machine learning classifier.

| Land type            | Location            | Area (m²) | Number of points |
|----------------------|---------------------|-----------|-----------------|
| Urban                | Madrid              | 14.1M     | 10.9M           |
|                      | Zamora              | 2.6M      | 2.7M            |
|                      | Salamanca           | 4.8M      | 3.9M            |
|                      | Navaleno            | 13.3M     | 18.9M           |
|                      | Vallejo de Valmedián| 10.4M     | 7.9M            |
| Non-urban (forest)   | Saelices del Payuelo| 9.8M      | 9.0M            |
| Non-urban (farmlands)| Sta. María del Páramo| 16.0M   | 11.6M           |

Figure 1. Fragment of aerial views of the training datasets: a) Madrid, b) Zamora, c) Salamanca, d) Navaleno, e) Vallejo de Valmedián, f) Saelices del Payuelo and g) Sta. María del Páramo.
Methodology

Methodologies based on the use of machine learning classifiers consist mainly of these three phases. A first phase of data ordering to generate the samples if the current data format is not valid as samples, named in this work as grid map generation. A second phase of calculation and extraction of useful features for the classifier. And a last phase where the features are used for the training and testing of the classifier.

Grid map generation

The structuring of a point cloud in a grid map consists of segmenting the cloud into regular squares of side \( l \) in the directions of X (longitude) and Y (latitude). Each square-cell contains enough number of points on which features can be extracted. This procedure involves a classification process, similar to a pixel-by-pixel classification (Weng, 2002), instead of a point-by-point classification (Wichmann et al., 2015). In this work, the benefits of this type of classification derive from the volume of data and land types considered for LOD0 classification.

When points are grouped in cells, the number of units to study is obviously reduced and features extraction for each cell on a grid map involves lower computational cost and lower time that the same process applied for each point. On the other hand, depending on cell size and land irregularity, this simplification can cause the existence of different land types in the same cell and consequently, a loss of accuracy. In cells that contains urban and non-urban points, the land type of the cell is considered as mode of land types of points on it.

Feature extraction

Once point cloud is structured in a grid map, geometric features are extracted for in each cell. More specifically, Z distribution and roughness are the features analysed in this work.

Z distribution is analysed through a histogram based on the distribution of the points in the Z direction. The histogram provides more information on the distribution of points along the Z value than common features such as maximum height, average height, variance and standard deviation (Golovinskiy et al., 2009). In addition, by normalizing the histogram to a percentage variation (passing the number of points to values between 0 and 1), a greater independence of density and number of points is achieved.

Roughness can be measured by different features (Weinmann et al., 2015; West et al., 2004): linearity \( L_3 \).

| Dataset                  | Area (m²) | Number of points | Urban points |
|-------------------------|-----------|------------------|--------------|
| Jerez de la Frontera    | 21.2M     | 13.1M            | 3.81M        |
| Zaragoza                | 9.2M      | 7.8M             | 3.85M        |
| Medina de Rioseco       | 11.5M     | 8.4M             | 0.85M        |

Figure 2. Fragment of aerial views of the testing datasets: a) Jerez de la Frontera, b) Zaragoza, c) Medina de Rioseco.
Classification is carried out through supervised machine learning techniques. This option has been chosen because it provides a high precision value in the final result (Nasrabadi, 2007). The use of a machine learning classifier needs an initial training phase, where patterns in the data are detected. A large amount of labelled and representative samples are required. The number of samples must be balanced, the same for each class (urban and non-urban). The efficiency of the classification is evaluated on a percentage of the training data, in this case. Six types of classifiers (Trees, Logistic Regression, SVM, SVM with Gaussian distribution, KNN and Ensemble Bagged Trees) and different parameters to feature extraction (grid size, point density and number of features) are used for classification. Their influence in the results are analysed in Section 5.1.

Results

Machine learning classifier analysis

In this section, the features that influence directly in the process of training and testing the machine learning detectors are analysed: classifier type, grid size, point cloud density and number of features.

Classifier type and data training

Six common classifiers are used in the training. They are based on Trees, Logistic Regression, SVM, SVM with Gaussian distribution, KNN and Ensemble Bagged Trees. Each dataset contributes randomly 1/8 of the total training samples, except the dataset of Madrid, that contributes 2/8 because it is the most extensive dataset and the only one that represents a big city, with a lot of geometric variations. A balanced number of samples between urban and non-urban are selected for training the classifier. Urban areas are divided equally into large cities (Madrid dataset) and small cities (Zamora and Salamanca datasets). Non-urban zones are divided equally into forest areas (Navaleno and Vallejo de Valmedián datasets) and agricultural zones (Saelices del Payuelo and Sta. María del Páramo). Each sample is represented by the 12 features explained in Section 3.2: five variables of the histogram related with the height and seven values of roughness (linearity, planarity, scattering, omnivariance, anisotropy, eigentropy and change of curvature). Within the training dataset, 75% is used for training and 25% for dev, selected to perform the accuracy of the detector.

Grid size

The optimum grid size is determined by an empirical analysis. A small grid size makes the classifier sensitive to local geometry variations, as isolated buildings. A large size involves an approximation error in the transition zones between urban and non-urban, where in a same cell, there are significant areas of transition zones between urban and non-urban, where in a same cell, there are significant areas of both land types. Figure 3 shows the accuracy variation according to grid size over the training dataset (trained with 500 samples with original density) from \( l = 20 \text{ m} \) to \( 500 \text{ m} \) (400 m\(^2\) to 250,000 m\(^2\)). The accuracy increases with the size of the grid, mainly because the training samples are not areas of transition. The relation between grid size and accuracy is evaluated on real case studies in Section 5.2, although to a minor extent.

Density

The density measured in the datasets is 1.3 points/m\(^2\), even though 0.5 points/m\(^2\) is the theoretical density provided by the datasheet of the “Instituto Geográfico Nacional”. However, there are density variations. In this section, the behaviour are of the classifier with the density decrease is evaluated. Figure 4 shows the variation in density when the samples are successively downsampled from original density to 5%, considering a grid size \( l = 100 \text{ m} \) (10,000 m\(^2\)) and with 500 training samples. Accuracy slightly decreases but it tends to remain stable in most classifiers. The detector works correctly even with low density (0.065 points/m\(^2\), 650 points per cell).
Features

Each feature has effects on the result in a different proportion. Characteristics with random distribution or characteristics equal to others do not provide any useful information for the training process. Figure 5 shows the box-plot of the features for urban land type and Figure 6 for non-urban land type. The Z distribution (Hist_Val_n) of the points into each cell are related with height of the elements: forest, farmlands and cities have different distributions. The first four values of the histogram bins present a distribution with somewhat higher values for urban than non-urban areas. The fifth value is the same in both lands types. The values of roughness features are clearly different by land types, except for the omnivariance in which values are extremely close to zero.

Figure 7 shows the contribution of combined features to the accuracy (tested in grid size of $l = 100$ m, original density, and with 500 samples). The use of only Z-distribution values contributes with 78.9% of accuracy in the classification. When adding one roughness feature such as the planarity, the accuracy is incremented to 88.5%.

As it can be seen in the difference between Figures 5 and 6, planarity has different distribution between land types, but there are also others that meet this condition. When using all features, the accuracy increases a 3.8% in comparison when using just one roughness feature. All roughness features are obtained from the same eigenvalues. With regard to computational cost for this example (grid size of $l = 100$ m with 500 samples), histogram feature extraction takes 0.23 s, histogram and planarity feature extraction takes 1.05 s and all histogram and roughness feature extraction takes 1.07 s. The greatest contributions are made by the histogram and one of the roughness features, but calculating them all spends 0.02 s more with an increment of accuracy of 3.8%.
Figure 5. Box-plot of feature value distribution for urban areas.

Figure 6. Box-plot of feature value distribution for non-urban areas.

Figure 7. Feature contribution to the accuracy.
Experimental study site results

In this section, four trained classifiers are evaluated on three real case studies with different grid sizes. They are selected among the results in training data, classifier type and grid size. All parameters are collected in Table 3. The results are shown in Tables 4 to 6. They have been accounted for based on the true urban, true non-urban, false urban, and false non-urban, the accuracy and the Cohen’s Kappa (Equation 8), which adjusts the effect of chance (Berry and Mielke, 1988; Cohen, 1960).

Figure 8 to 10 shows the results of the classification for grid size \(l = 500\) m.

\[
k = \frac{P_o - P_e}{1 - P_e} \tag{8}
\]

Where \(P_o = \sum P_{ui}\) (accuracy of each land type) and \(P_e = \sum P_i \cdot P_i\) (probability at the same response for each land type).

Tables 4 to 6 show that classification is performed with an accuracy around 90% in all tests except for the case study 1 where accuracy falls down to 68%. The Cohen’s Kappa index strongly varies for case studies: 0.68 for case study 1, 0.78 for case study 2 and 0.47 for case study 3. The grid size \(l = 500\) m presents strong variations with very good results in case study 2 \((k = 0.81)\) and very bad in the other two cases \((k = 0.22\) and \(k = 0.31)\). Many misclassifications are concentrated in the areas of transition between urban and non-urban, even in some isolated buildings or small concentration in non-urban areas (Figures 8 and 10). The datasets for training are integrated by zones of each type. They do not contain transition zones. The response of the classifier in transition zones is defined by the most relevant features in each cell. This can be seen more clearly with a larger grid size: \(l = 500\) m (Figure 11).

In the case study 2 (Figure 9), there are areas of false urban positives on riversides. The forests used for training are forests in mountainous areas while the forest of the case study is a fluvial forest. Each type of forest presents specific characteristics that can be very varied among themselves (Antonarakis, Richards, & Brasington, 2008). As forests in fluvial zones have not been used in the training of the classifier, this type of land use concentrates false positives. An example of this can be seen along the river in case study 2 and along the small rivers in case study 3 (Figure 10).

There are some constructive and urbanization differences between different cities. The urban areas used for the training are from the center of Spain while the case study 1 is located in the South. This affects the features of urban areas, although to a small extent. For example, case study 1 (Figures 8 and 11) presents concrete areas that do not exist in the training samples, as transport and chalets areas. Consequently, they are detected as false non-urban positives.

At last, processing times are collected in Tables 7 to 9. The algorithm has been implemented in Matlab and processed on an Intel Core i7-7700HQ CPU 2.80 GHz with 16GB RAM. The processing time is divided into three parts: grid generation, feature extraction and the classification. The grid

---

Table 3. Classifiers employed on real case studies.

| Type          | Grid size | Density     | Number of features | Number of training samples |
|---------------|-----------|-------------|--------------------|----------------------------|
| SVM           | 50 m      | 0.5 points/m² | 12                | 1000                       |
| Ensemble      | 100 m     | 0.5 points/m² | 12                | 1000                       |
| Bagged Trees  | 200 m     | 0.5 points/m² | 12                | 1000                       |
| Tree          | 500 m     | 0.5 points/m² | 12                | 1000                       |
| KNN           |           |             |                   |                            |

Table 4. Results of the classification in case study 1.

| Grid size (m) | T urban | T non-urban | F urban | F non-urban | Accuracy | Kappa |
|---------------|---------|-------------|---------|-------------|----------|-------|
| 50            | 1410    | 6100        | 93      | 977         | 0.88     | 0.65  |
| 100           | 453     | 1535        | 15      | 162         | 0.92     | 0.78  |
| 200           | 95      | 389         | 14      | 62          | 0.86     | 0.63  |
| 500           | 11      | 49          | 12      | 16          | 0.68     | 0.22  |

Table 5. Results of the classification in case study 2.

| Grid size (m) | T urban | T non-urban | F urban | F non-urban | Accuracy | Kappa |
|---------------|---------|-------------|---------|-------------|----------|-------|
| 50            | 1490    | 1767        | 339     | 128         | 0.87     | 0.75  |
| 100           | 383     | 462         | 75      | 16          | 0.90     | 0.80  |
| 200           | 87      | 121         | 14      | 12          | 0.89     | 0.77  |
| 500           | 16      | 24          | 2       | 2           | 0.91     | 0.81  |

Table 6. Results of the classification in case study 3.

| Grid size (m) | T urban | T non-urban | F urban | F non-urban | Accuracy | Kappa |
|---------------|---------|-------------|---------|-------------|----------|-------|
| 50            | 274     | 4033        | 365     | 200         | 0.88     | 0.43  |
| 100           | 89      | 1019        | 81      | 29          | 0.91     | 0.57  |
| 200           | 15      | 268         | 18      | 14          | 0.90     | 0.43  |
| 500           | 1       | 49          | 0       | 4           | 0.93     | 0.31  |
generation depends on the size of the dataset, there are no significant differences between grid sizes. Time used in feature extraction inversely proportional to the grid size, for the same dataset, a low grid size involves a larger number of cells (higher number of iterations) with fewer points and more computing time, by contrast, a high grid size involves a low number of cells (lower number of

| Model type | Grid size (m) | Grid generation (s) | Feature extraction (s) | Classification (s) | Total (s) |
|-----------|---------------|---------------------|------------------------|--------------------|-----------|
| SVM       | 50            | 6.98                | 23.62                  | 1.14               | 31.74     |
| Ensemble Bagged Trees | 100            | 6.53                | 8.58                   | 0.32               | 15.43     |
| Tree      | 200           | 6.34                | 4.71                   | 0.08               | 11.12     |
| KNN       | 500           | 6.25                | 3.40                   | 0.68               | 10.34     |
iterations) with more number of points. The classification time depends on the type of classification model; models based on trees need less time to classify than the others.

Conclusions

A methodology has been designed to classify ALS point clouds in urban and non-urban areas, giving solution to an LOD0 defined in point clouds for later use according to the desired work. The difference between both is related to concentration of buildings and infrastructures. The methodology consists in a grid map generation from airborne point clouds, a feature extraction and a machine learning classification. Grid map generation structures point cloud in cells of a regular size on the XY plane. For each cell of the grid map, features related to Z distribution (histogram of five bins) and roughness (linearity, planarity, scattering, eigentropy, change of curvature, omnivariance and anisotropy) are extracted. Six machine learning classifiers have been trained with samples from seven areas, including cities, forests and farmlands. The methodology has been tested on three real case studies.

The results in real cases show a high accuracy (around 90%), tested in different grid sizes and have an optimal accuracy with a grid size $l = 100$ m. The Cohen’s Kappa index varies according to the characteristics of the case studies, reaching 0.81 in the best case. The misclassifications have been detected mainly in transition zones (cells with urban and non-urban points), areas with fluvial vegetation, housing areas and concentration of means of transport, all of which are not contemplated in the training datasets. The processing time is very fast, large areas of land (millions of points) are classified in a few seconds.

In summary, the designed methodology obtains very good results for a LOD0 classification of urban and non-urban areas, using exclusively the geometry provided by the ALS point clouds, without requiring any type of colour or laser intensity information employed by other methods reported in the literature. Its short processing time and robustness against density changes make its use feasible in the classification of large-scale point clouds. Density analyses seem to indicate that the increase in accuracy is related to an increase in density. In the future, it is interesting to assess the extent to which this relationship is at its maximum, as well as to evaluate the time and economic costs involved in this relationship in terms of acquisition, storage and processing of higher density data. Also for future improvements, different features such as RGB, intensity and number of returns will be evaluated with the purpose of achieve the highest possible quality in the classification.

Also, Deep Learning techniques with neural networks will be implemented.

Acknowledgments

Authors would like to thank to Xunta de Galicia for the financial support given through human resources grant (ED481B 2016/079-0) and competitive reference groups (ED431C 2016-038), and the Ministerio de Economía, Industria y Competitividad -Gobierno de España-(TIN2016-77158-C4-2-R, RTC-2016-5257-7). Also, authors would like to thank to Universidade de Vigo for the financial support (00VI 131H 641.02) and this work has been partially supported by the Dirección General de Tráfico (Spanish Ministry of Interior) (Grant SPIP2017-02122). The statements made herein are solely the responsibility of the authors.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Xunta de Galicia [ED431C 2016-038,ED481B 2016/079-0], Ministerio de Economía, Industria y Competitividad -Gobierno de España [RTC-2016-5257-7,TIN2016-77158-C4-2-R];Dirección General de Tráfico (Spanish Ministry of Interior) [SPIP2017-02122],Universidade de Vigo [00VI 131H 641.02].

| Table 8. Processing time in classification of case study 2. |
| --- |
| Model type | Grid size (m) | Grid generation (s) | Feature extraction (s) | Classification (s) | Total (s) |
| SVM | 50 | 3.60 | 8.58 | 0.56 | 12.94 |
| Ensemble Bagged Trees | 100 | 3.50 | 4.49 | 1.07 | 9.07 |
| Tree | 200 | 3.50 | 2.45 | 0.03 | 5.98 |
| KNN | 500 | 3.58 | 2.02 | 0.09 | 5.69 |

| Table 9. Processing time in classification of case study 3. |
| --- |
| Model type | Grid size (m) | Grid generation (s) | Feature extraction (s) | Classification (s) | Total (s) |
| SVM | 50 | 4.31 | 14.59 | 0.62 | 19.53 |
| Ensemble Bagged Trees | 100 | 4.35 | 6.69 | 0.62 | 11.66 |
| Tree | 200 | 3.88 | 2.86 | 0.06 | 6.81 |
| KNN | 500 | 3.91 | 2.26 | 0.48 | 6.66 |
ORCID

J. Balado  http://orcid.org/0000-0002-3758-3102
L. Díaz-Vilarino  http://orcid.org/0000-0002-2382-9431
L. M. González-Desantos  http://orcid.org/0000-0002-3898-7854

References

Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R.E., & Peck, D.L. (1976). A land use and land cover classification system for use with remote sensor data. US Gov Print Office, 964, 41.

Antonarakis, A.S., Richards, K.S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. Remote Sens Environment, 112, 2988–2998.

Balado, J., Díaz-Vilarino, L., Arias, P., & González-Jorge, H. (2018). Automatic classification of urban ground elements from mobile laser scanning data. Automation in Construction, 86, 226–239.

Berry, K.J., & Mielke, P.W., Jr. (1988). A generalization of Cohen’s Kappa agreement measure to interval measurement and multiple raters. Education Psychologist Measurement, 48, 921–933.

Biljecki, F., Ledoux, H., & Stoter, J. (2016). An improved LOD specification for 3D building models. Computation Environment Urban Systems, 59, 25–37.

Brennan, R., & Webster, T.L. (2006). Object-oriented land cover classification of lidar-derived surfaces. Canada Journal Remote Sensing, 32, 162–172.

Cao, Y., Wei, H., Zhao, H., & Li, N. (2012). An effective approach for land-cover classification from airborne lidar fused with co-registered data. International Journal Remote Sensing, 33, 5927–5935.

Castelluccio, M., Poggi, G., Sansone, C., & Verdoliva, L. (2015). Land use classification in remote sensing images by convolutional neural networks. arXiv preprint arXiv:1508.00092.

Census Bureau United States, 2015. 2010 census urban and rural classification and urban area criteria [WWW document]. Retrieved www.census.gov/geo/reference/ua/urban-rural-2010.html from January 19, 2018.

Centro Nacional de Información Geográfica, n.d. Centro de descargas [WWW Document]. URL www.centrode descargas.cni.ge (accessed January 19, 2018).

Cohen, J. (1960). A coefficient of agreement for nominal scales. Education Psychologist Measurement, 20, 37–46.

Department For Environment Food & Rural Affairs, n.d. Environmental open-data applications and datasets [WWW Document]. Retrieved environment.data.gov.uk from June 18, 2018.

Fu, P., & Weng, Q. (2016). A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. Remote Sensing Environment, 175, 205–214.

Gamba, P., & Houshmand, B. (2002). Joint analysis of SAR, LIDAR and aerial imagery for simultaneous extraction of land cover, DTM and 3D shape of buildings. International Journal Remote Sensing, 23, 4439–4450.

Gerke, M., & Xiao, J. (2014). Fusion of airborne laserscanning point clouds and images for supervised and unsupervised scene classification. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 78–92.

Golovinskiy, A., Kim, V.G., & Funkhouser, T. (2009). Shape-based recognition of 3d point clouds in urban environments. IEEE 12th International Conference on Computer Vision, 2009, 2154–2161.

Guan, H., Ji, Z., Zhong, L., Li, J., & Ren, Q. (2013). Partially supervised hierarchical classification for urban features from lidar data with aerial imagery. International Journal Remote Sensing, 34, 190–210.

Guo, L., Chehata, N., Mallet, C., & Boukir, S. (2011). Relevance of airborne lidar and multispectral image data for urban scene classification using random forests. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 56–66.

Hartfield, K.A., Landau, K.L., & van Leeuwen, W.J.D. (2011). Fusion of high resolution aerial multispectral and LiDAR data: Land cover in the context of urban mosquito habitat. Remote Sensing. doi:10.3390/rs3112364

Huang, H., Chen, Y., Clinton, N., Wang, J., Wang, X., Liu, C., … Zhu, Z. (2017). Mapping major land cover dynamics in Beijing using all Landsat images in Google earth engine. Remote Sens Environment, 202, 166–176.

Huang, M.-J., Shyue, S.-W., Lee, L.-H., & Kao, -C.-C. (2008). A knowledge-based approach to urban feature classification using aerial imagery with lidar data. Photogramm Engineering Remote Sensing. doi:10.14358/ PERS.74.12.1473

Hui, L., Di, L., Xianfeng, H., & Deren, L., 2008. Laser intensity used in classification of lidar point cloud data, in: IGRASS 2008–2008 IEEE International Geoscience and Remote Sensing Symposium. p. II-1140-II-1143. doi:10.1109/IGARSS.2008.4779201

Jebur, M.N., Mohd Shafri, H.Z., Pradhan, B., & Tehrany, M.S. (2014). Per-pixel and object-oriented classification methods for mapping urban land cover extraction using SPOT 5 imagery. Geocarto International, 29, 792–806.

Kumar, P., Lewis, P., & McCarthy, T. (2017). The potential of active contour models in extracting road edges from mobile laser scanning data. Infrastructures. doi:10.3390/ infrastructures2030009

Lague, D., Brodou, N., & Leroux, J. (2013). Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). ISPRS Journal of Photogrammetry and Remote Sensing, 82, 10–26.

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery. Remote Sensing. doi:10.3390/rs6020964

Licciardi, G., & Chanussot, J. (2018). Spectral transformation based on nonlinear principal component analysis for dimensionality reduction of hyperspectral images. European Journal Remote Sensing, 51, 375–390.

Lu, D., & Weng, Q. (2006). Use of impervious surface in urban land-use classification. Remote Sens Environment, 102, 146–160.

Luo, S., Wang, C., Xi, X., Zeng, H., Li, D., Xia, S., & Wang, P. (2016). Fusion of airborne discrete-return LiDAR and hyperspectral data for land cover classification. Remote Sensing. doi:10.3390/rs8010003

Momeni, R., Aplin, P., & Boyd, D.S. (2016). Mapping complex urban land cover from spaceborne imagery: The influence of spatial resolution, spectral band set and classification approach. Remote Sensing, 8. doi:10.3390/rs8020088

Moreira, S., Fernandez, P., Pereira, L., & Moreira, M. (2017). Improving a DSM obtained by unmanned aerial vehicles for flood modelling. IOP Conference Series Earth Environment Sciences, 95, 22014.
Morsy, S., Shaker, A., El-Rabbany, A., & LaRocque, P.E. (2016). Airborne multispectral Lidar data for land-cover classification and land/water mapping using different spectral indices. *ISPRS Annals Photogrammetry Remote Sensing and Spatial Information Sciences* 3-3, 217–224. doi:10.5194/isprs-annals-III-3-217-2016

Nacional, I.G., n.d. Instituto Geográfico Nacional [WWW Document]. Retrieved www.ign.es from August 28, 2018.

Nasrabadi, N.M. (2007). Pattern recognition and machine learning. Journal of electronic imaging, 16, 49901.

Netherlands eScience Center, n.d. AHN2 [WWW Document]. Retrieved ahn2.pointclouds.nl from June 18, 2018.

Pang, S., Hu, X., Wang, Z., & Lu, Y. (2014). Object-based analysis of airborne LiDAR data for building change detection. *Remote Sensing*, 6, 10733–10749.

Piedelobo, L., Ortega-Terol, D., Del Pozo, S., Hernández-López, D., Ballesteros, R., Moreno, M.A., ... González-Agulera, D. (2018). HidroMap: A new tool for irrigation monitoring and management using free satellite imagery. *ISPRS International Journal Geo-Information*, 7. doi:10.3390/ijgi7060220

Poursanidis, D., Chrysoulakis, N., & Mitraka, Z. (2015). Landsat 8 vs. Landsat 5: A comparison based on urban and peri-urban land cover mapping. *International Journal of Applied Earth Observation and Geoinformatics*, 35, 259–269.

Priestnall, G., Jaafar, J., & Duncan, A. (2000). Extracting urban features from LiDAR digital surface models. *Computation Environment Urban Systems*, 24, 65–78.

Qian, Y., Zhou, W., Yan, J., Li, W., & Han, L. (2015). Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sensing*. doi:10.3390/rs7100153

Rottersteiner, F., & Briese, C. (2002). A new method for building extraction in urban areas from high-resolution LiDAR data. *International Archives Photogrammetry Remote Sensing*, XXXIV, 295–301. doi:10.1051/s12056347

Saadat, H., Adamowski, J., Bonnell, R., Sharifi, F., Namdar, M., & Ale-Ebrahim, S. (2011). Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 608–619.

Salah, M., Trinder, J., & Shaker, A. (2009). Evaluation of the self-organizing map classifier for building detection from lidar data and multispectral aerial images. *Journal of Photogrammetry and Remote Sensing*, 54, 15–34.

Scaini, M., Höfle, B., Baugarten Kersting, A.P., Barazzetti, L., Previtali, M., & Wujanz, D. (2018). Methods from information extraction from Lidar intensity data and multispectral Lidar technology. *ISPRS - International Archives Photogrammetry Remote Sensing and Spatial Information Sciences*, XXII-3, 1503–1510. doi:10.5194/isprs-archives-XLI-III-1503-2018

Schmidheiny, K., & Suedekum, J. (2015). The pan-European population distribution across consistently defined functional urban areas. *Economics Letters*, 133, 10–13.

Serna, A., & Marcotegui, B. (2013). Urban accessibility diagnosis from mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 84, 23–32.

Serna, A., Marcotegui, B., & Hernández, J. (2016). Segmentation of façades from urban 3D point clouds using geometrical and morphological attribute-based operators. *ISPRS International Journal Geo-Information*, 5. doi:10.3390/ijgi5010006

Stow, D., Lopez, A., Lippitt, C., Hinton, S., & Weeks, J. (2007). Object-based classification of residential land use within Accra, Ghana based on QuickBird satellite data. *International Journal Remote Sensing*, 28, 5167–5173.

Teo, T.-A., & Shih, T.-Y. (2013). Lidar-based change detection and change-type determination in urban areas. *International Journal Remote Sensing*, 34, 968–981.

Teo, T.-A., & Wu, H.-M. (2017). Analysis of land cover classification using multi-wavelength LiDAR system. *Applications Sciences*, 7. doi:10.3390/app7070663

United Nations Statistics Division. (2006). Demographic Yearbook. 2005.

Vosselman, G., & Liang, Z. (2009). Detection of curbstones in airborne laser scanning data. *Geoinf Sciences*, 38, 293–304.

Weinmann, M., Zutzi, B., Hinz, S., & Mallet, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 286–304.

Weng, Q. (2002). Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling. *Journal of Environmental Management*, 64, 273–284.

West, K.F., Webb, B.N., Lersch, I.R., Pothier, S., Triscari, J. M., & Iverson, A.E., 2004. Context-driven automated target detection in 3D data, in: Proceedings of SPIE 5426, Automatic Target Recognition XIV, SPIE, Orlando, USA, pp. 133–143. doi:10.1117/12.542536

Wichmann, V., Bremer, M., Lindenberger, J., Rutzinger, M., Christian, G., & Frederic, P.-M. (2015). Evaluating the potential of multispectral airborne Lidar for topographic mapping and land cover classification, ISPRS annals of photogrammetry. *Remote Sensing and Spatial Information Sciences*. doi:10.5194/isprsannals-III-3-W5-113-2015

Yan, W.Y., Shaker, A., & El-Ashmawy, N. (2015). Urban land cover classification using airborne LiDAR data: A review. *Remote Sens Environment*, 158, 295–310.

Yu, B., Liu, H., Zhang, L., & Wu, J., 2009. An object-based two-stage method for a detailed classification of urban landscape components by integrating airborne LiDAR and color infrared image data: A case study of downtown Houston, in: 2009 Joint Urban Remote Sensing Event. pp. 1–8. doi:10.1109/URS.2009.5137543

Zarea, A., & Mohammadzadeh, A. (2016). A novel building and tree detection method from LiDAR data and aerial images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9, 1864–1875.

Zhang, J., & Lin, X. (2017). Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *International Journal Image Data Fusion*, 8, 1–31.

Zhang, Q., Qin, R., Huang, X., Fang, Y., & Liu, L. (2015). Classification of ultra-high resolution orthophotos combined with DSM using a dual morphological top hat profile. *Remote Sensing*. doi:10.3390/rs71215840

Zou, X., Zhao, G., Li, J., Yang, Y., & Fang, Y. (2016). 3D land cover classification based on multispectral Lidar point clouds. *ISPRS - International Archives Photogrammetry Remote Sensing and Spatial Information Sciences*, XLI-B1, 741–747. doi:10.5194/isprs-archives-XLI-B1-741-2016