Classification of Vegetation in Aerial LiDAR Data

Denis Horvat
Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia
E-mail: denis.horvat@um.si, Web: https://gemma.feri.um.si/

Thesis summary

Keywords: LiDAR, classification, vegetation, mathematical morphology, surface fitting

Received: April 12, 2017

This contribution summarises a doctoral dissertation which proposes an algorithm for the classification of vegetation points in aerial LiDAR data. The algorithm characterizes vegetated areas based on statistically large dispersion in elevations of points, and the context in which the points are located. The algorithm is able to classify vegetation in both rural and urban areas with an average F1 score of 97.9% and 91.0%, respectively. The point-clouds can contain different types of vegetation and various degrees of canopy densities.

Povzetek: V predlaganem prispevku povzamemo doktorsko disertacijo, ki predlaga algoritem za klasifikacijo točk vegetacije iz podatkov LiDAR. Algoritem ovrednoti območja vegetacije na podlagi statistično visoke razpršenosti višin točk v kombinacji s kontekstom v katerem se točke nahajajo. Algoritem klasificira točke vegetacije v urbanih in neurbanih področjih s 97% in 91% povprečnim rezultatom F1. Oblaki točk lahko vsebujejo različne tipe vegetetacije z različno gostoto olistanosti.

1 Introduction

The potential of the data obtained by the aerial LiDAR (**Li**ght **D**etection **a**nd **R**anging) systems has been utilised increasingly by a variety of scientific and industrial applications. Data acquisition using LiDAR produces a point-cloud, in which the individual points are calculated based on the time delay between the emitted and detected laser beam. While the obtained point-clouds from a plane-mounted LiDAR can represent the underlying Earth's surface accurately, the entity to which a given point belongs (e.g. ground, building, vegetation) is not known. This contextual knowledge is crucial for a variety of applications, such as environmental simulations, urban planning, or the generation of a canopy height model. Thus, preprocessing, using a classification algorithm is usually applied, in which each point is correlated with one of the predetermined classes.

This paper is a summary of a PhD thesis [1] (and the corresponding paper [4]), which proposed an algorithm for classification of vegetation points within LiDAR data. Vegetation can be particularly challenging to identify, as the classification model should be universal enough to cover the various sizes, shapes and canopy densities of different vegetation but, at the same time, still differentiate it from other surface objects (e.g. houses, cars, fences). The following section outlines the proposed classification algorithm, while section 3 evaluates it. Section 4 concludes the paper.

2 Classification algorithm

Clusters of points that represent vegetation are, in most cases, defined by statistically large dispersions in elevation. This is caused by the vegetation's non-linear shape and porosity, as the laser beam can usually penetrate the canopy and capture many points within, or even under, the vegetation. The mentioned properties can be characterized efficiently by modifying the LoFS (**Lo**cal **F**itting of **S**urfaces) method [2]. Namely, by locally fitting planes on the LiDAR-derived surface and evaluating the fitting error using all points in the fitted area, a distinction can be made between vegetation and most of other man-made objects. Larger fitting errors are expected in the former case, while the latter ones usually produce errors that are identifiably smaller.

The remaining non-vegetation that does not conform to this definition (i.e. also produces a larger fitting error) is handled using contextual analysis which defines: 1) Attached objects 2) Overgrowing vegetation and 3) Small objects. Attached objects represents a transition between areas (e.g. a wall, balcony or chimney). Such objects can be identified (and subsequently removed) using spatially-variant morphological dilation [3] where all non-ground areas with small fitting errors are dilated. The extent of the dilation is controlled locally, using a structuring element with the radius equal to the distance from the nearest ground. Spatially-variant dilation is used similarly on areas with high fitting errors to identify overgrown vegetation. However, the radius, in this case, is dependent on the height difference between a given area and the nearest non-ground area with a small fitting error. Lastly, small objects are removed using connected operators. The final result that includes the described fitting error evaluation and contextual analysis is then mapped onto the individual LiDAR points to get the classification.

3 Results

The algorithm was tested on multiple rural and urban datasets which contained different types of vegetation and degrees of canopy densities. The results were evaluated by counting the false positives, false negatives, true positives and true negatives which served as an input for the calculation of the well-established F1 score. An average F1 score of 97.9% was achieved for rural and 91.0% for urban environments which, because of the more complex geometry, tend to be more difficult to classify. The use of contextual analysis improves the results in urban areas significantly. Namely, by removing the attached object, small object and handling overgrown vegetation, the F1 score is improved, on average, by 8.8%, 1.1% and 1.0%, respectively.

4 Conclusion

The PhD thesis presented an algorithm for the classification of LiDAR data, which classifies vegetated areas with high accuracy in characteristically different point-clouds (rural environment, urban environment, different types of vegetation, various leaf-on conditions). Additionally, the algorithm can be used in most classification scenarios that contain aerial point-clouds, as it relies only on geometrical features of the point-cloud and the subsequent contextual analysis.

Acknowledgement

The authors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P2-0041 and project J2-6764)

References

[1] D. Horvat (2017) Algorithm for classification of vegetation in LiDAR data, *Doctoral dissertation, Faculty of Electrical Engineering and Computer Science, University of Maribor (in Slovene).*

[2] D. Mongus, N. Lukač and B. Žalik (2014) Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 93, pp. 145–156.

[3] N. Bouaynaya and Mohammed Charif-Chefchaouni (2008) Theoretical Foundations of Spatially-Variant Mathematical Morphology Part I: Binary Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 5, pp. 145–156.

[4] D. Horvat, B. Žalik and D. Mongus (2016) Context-dependent detection of non-linearly distributed points for vegetation classification in airborne LiDAR, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 116, pp. 1–14.