ISPRS Journal of Photogrammetry and Remote Sensing 137 (2018) 73–83

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

3D micro-mapping: Towards assessing the quality of crowdsourcing to support 3D point cloud analysis

Benjamin Herfort *, Bernhard Höfle, Carolin Klonner

GIScience Research Group, Department of Geography, Heidelberg University, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 2 August 2017
Received in revised form 27 December 2017
Accepted 10 January 2018

Keywords:
LiDAR
Urban trees
Crowdsourcing
Point cloud classification
Quality

ABSTRACT

In this paper, we propose a method to crowdsource the task of complex three-dimensional information extraction from 3D point clouds. We design web-based 3D micro tasks tailored to assess segmented LiDAR point clouds of urban trees and investigate the quality of the approach in an empirical user study. Our results for three different experiments with increasing complexity indicate that a single crowdsourcing task can be solved in a very short time of less than five seconds on average. Furthermore, the results of our empirical case study reveal that the accuracy, sensitivity and precision of 3D crowdsourcing are high for most information extraction problems. For our first experiment (binary classification with single answer) we obtain an accuracy of 91%, a sensitivity of 95% and a precision of 92%. For the more complex tasks of the second Experiment 2 (multiple answer classification) the accuracy ranges from 65% to 99% depending on the label class. Regarding the third experiment – the determination of the crown base height of individual trees – our study highlights that crowdsourcing can be a tool to obtain values with even higher accuracy in comparison to an automated computer-based approach. Finally, we found out that the accuracy of the crowdsourced results for all experiments is hardly influenced by characteristics of the input point cloud data and of the users. Importantly, the results' accuracy can be estimated using agreement among volunteers as an intrinsic indicator, which makes a broad application of 3D micro-mapping very promising.

1. Introduction

The automated detection and classification of objects from 3D geodata (e.g. derived via LiDAR or photogrammetry) is an important part in many geographic tasks such as geomorphological mapping (Bremer and Sass, 2012; Höfle and Rutzinger, 2011; Hohenthal et al., 2011), the generation of 3D city models (Golovinskiy et al., 2009; Niemeyer et al., 2012) or in robotics and autonomous driving (Levinson et al., 2011; Maturana and Scherer, 2015). Especially the computer-based detection and classification of single objects in urban environments still face many challenges due to complex object structures and a wide range of different semantic classes in close neighborhoods. Examples for such complex structures are buildings, roads, urban furniture, trees, and a multitude of temporary objects such as persons and cars (Niemeyer et al., 2012).

Regarding automated computer-based methods for object detection, data created by humans often play an essential role during the very initial steps of an analysis conducted such as for generating independent training and validation data. For instance, Niemeyer et al. (2012) investigate the performance of an automated airborne LiDAR point cloud classification using geometric features and intensity values to classify points into building, low vegetation, tree, natural ground and asphalt. Douillard et al. (2011) and Koenig and Höfle (2016) underline the importance of segmentation as a pre-processing task in several autonomous perception problems. Similarly, artificial neural networks have shown remarkable capabilities to process information and solve problems by learning from examples in different research domains. Additionally, most of the studies apply several pre-processing mechanisms in which the initial steps consist of general data inspection, noise removal or deletion of outliers; mostly done manually (Basheer and Hajmeer, 2000). However, as training and reference datasets, and during the pre-processing, all authors used several sets of manually labelled data to find the best parameter settings for their automated approaches.

* Corresponding author.
E-mail addresses: herfort@uni-heidelberg.de (B. Herfort), hoefle@uni-heidelberg.de (B. Höfle), C.Klonner@uni-heidelberg.de (C. Klonner).

https://doi.org/10.1016/j.isprsjprs.2018.01.009
0924-2716/

While the automated methods are getting more powerful, manual pre-processing or the evaluation of the derived methods remain bottlenecks, since they still depend on humans analyzing the data and cannot be done automatically with sufficient accuracy. This is time consuming for some tasks, and thus it might not be a feasible solution for other tasks which incorporate larger datasets. In addition, manual annotation by single people might be subjective and can lead to an operator bias, which is problematic for the benchmark of new algorithms and tools.

Crowdsourcing can be seen as a complementing approach for such complex information extraction. In general, crowdsourcing campaigns face three different challenges (Barrington et al., 2011): (1) How can you split the overall task into single components (micro tasks)? (2) How can you organize a large group of people to contribute? (3) How can you aggregate the individual answers into a coherent overall result?

There are already many crowdsourcing tasks, which focus on information extraction for two-dimensional data such as photographs or satellite imagery. In computer vision, Peekaboom was one of the first large scale attempts to crowdsource information valuable for computer vision algorithms (von Ahn et al., 2006). Russakovsky et al. (2015) present a system that collects bounding box annotations through crowdsourcing using imageNet data for learning object detectors, whereas Deng et al. (2013) introduced an online game to crowdsource discriminative features from photographs to train a classifier to distinguish bird types.

In the domain of geospatial information science, the detection of damaged or destroyed buildings from satellite imagery is another prominent crowdsourcing task (Barrington et al., 2011; Kerle and Hoffman, 2013; Westrope et al., 2014). Herfort et al. (2016) reveal that there are three types of crowdsourcing geographic information from satellite imagery which pose an increasing complexity to the users: classification, digitization and conflation. Regarding the mapping of vegetation, the Forest Watchers citizen science project allows volunteers around the globe to review satellite images of forested regions and confirm whether automatic assignments of forested and deforested regions are correct (Arcanjo et al., 2016).

Despite this focus on crowdsourcing tasks, which incorporate two-dimensional data in various disciplines, we are predominantly facing three-dimensional recognition problems in our daily lives. However, the human brain is exceptionally suited for solving such a kind of problems. Furthermore, we also get used to the digital visualization of three-dimensional geodata on 2D screens and displays, which is becoming increasingly prevalent in many domains such as the movie or gaming industries. However, so far, only little work has been done regarding the 3D environments (Goetz and Zipf, 2013).

Hara et al. (2014) present an approach that combines machine learning techniques, computer vision and crowdsourcing to extract 2.5D information on curb ramps from Google street view images. The results of this study highlight that the combined approach can achieve results with the same quality as manually labelled datasets, however, their approach is able to reduce the time costs by 13%.

However, most crowdsourcing projects still rely on two-dimensional datasets as an input and only the processed output contains a three-dimensional representation in the data. To the best of our knowledge there has been no crowdsourcing project so far, where users interact directly with 3D point cloud data. Furthermore, so far, crowdsourcing has not been explored regarding the annotation of vegetation objects in an urban environment. Therefore, there is a lack of understanding how the methods already applied to other domains can be used and transferred to the even more complex 3D urban environments and vegetation objects.

The subsequent assumption of our study is that object detection and classification in LiDAR point cloud data are tasks to which crowdsourcing can add a great beneficial value. On the one hand, the automated detection and classification of these single parts of the point cloud still faces many challenges. It has been shown that the detection and classification of trees in an urban environment is difficult to solve fully automatically in comparison to other objects such as buildings (Niemeyer et al., 2012). On the other hand, many processing workflows rely on a prior segmentation of the point cloud data into individual clusters of points which are spatially associated (Koenig and Höfle, 2016). The results of this segmentation step can function as an ideal basis for the definition of the crowdsourcing tasks.

Therefore, we are designing and investigating 3D micro tasks tailored to complex three-dimensional information extraction problems, which cannot be solved automatically with high accuracy. The crowdsourced answers are examined in an empirical user study and are analyzed regarding the accuracy of the information derived with respect to different factors related to the input data and user characteristics. Hence, this paper addresses the following main research question: How do characteristics of point cloud data and user characteristics affect the quality of the results of three-dimensional crowdsourcing tasks? Subsequently, we need to answer two important sub-questions: (a) What is the accuracy, specificity and precision of three-dimensional crowdsourcing point cloud classification tasks when aggregating with majority answer? (b) What is the accuracy of three-dimensional crowdsourcing tasks to assess the crown base height of a tree in comparison to the automatically extracted height?

These findings are fundamental for enhancing our understanding of how 3D crowdsourcing works in order to learn to design new future tasks and to find out about chances and limitations of this innovative branch of crowdsourcing.

2. Methods

Crowdsourcing tends to be promising when the information extraction problems tackle different levels of complexity for both human interpreters and automated algorithms. Hence, an ideal crowdsourcing project is characterized by the intermingling of different types of information that needs to be recognized. We designed three experiments to address the research questions and each is explained in detail in the following section and addresses a different kind of information extraction problem. In these experiments the volunteers will work on several so-called Micro-Mapping tasks which deal with the same information extraction problem. We will evaluate our approach by an extensive case study and by bringing together the results of the three experiments.

2.1. Conceptual approach

We concentrate on the crowdsourced assessment of segmented airborne LiDAR point clouds representing urban trees. The types of information which can be extracted are various and cover (1) the detection of objects superimposing a tree (e.g. cars), (2) the detection of segmentation errors caused by the automatic pre-processing algorithms (e.g. dissected tree crowns) and (3) the detection of parts of the tree that are not depicted in the point cloud (e.g. missing trunk due to low point density). An additional problem to consider regarding trees in general is (4) the detection of the crown base, which is an important parameter to further differentiate individual tree species or reconstruct forest stands (Koma et al., 2016; Morsdorf et al., 2004; Yu et al., 2014).