AN ADAPTED CONNECTED COMPONENT LABELING FOR CLUSTERING NON-PLANAR OBJECTS FROM AIRBORNE LIDAR POINT CLOUD

B. Aissou^{1,*}, A. Belhadj Aissa¹

¹ University of Sciences and Technology of Algiers, Image Processing and Radiation Laboratory, Bab Ezzouar, 16111, Algeria (bahae.aissou, houria.belhadjaissa)@gmail.com

Commission II, WG II/3

KEY WORDS: LiDAR, Connected component analysis, Clustering, Non-planar surfaces, Dormers, Chimneys

ABSTRACT:

Light Detection And Ranging (LiDAR) is an active remote sensing technology used for several applications. A segmentation of Airborne Laser Scanning (ALS) point cloud is very important task that still interest many scientists. In this paper, the Connected Component Analysis (CCA), or Connected Component Labeling is proposed for clustering non-planar objects from Airborne Laser Scanning (ALS) LiDAR point cloud. From raw point cloud, sub-surface segmentation method is applied as preliminary filter to remove planar surfaces. Starting from unassigned points, CCA is applied on 3D data considering only neighboring distance as initial parameter. To evaluate the clustering, an interactive labeling of the resulting components is performed. Then, components are classified using Support Vector Machine, Random Forest and Decision Tree. The ALS data used is characterized by a low density (4-6 points/m²), and is covering an urban area, located in residential parts of Vaihingen city in southern Germany. The visualization of the results shown the potential of the proposed method to identify dormers, chimneys and ground class.

1. INTRODUCTION

Light Detection And Ranging (LiDAR) is an active remote sensing technology used for several applications. A segmentation of Airborne Laser Scanning (ALS) point cloud is very important task that still interest many scientists. Most of point clouds segmentation methods are typically adapted to extract planar surfaces as roof faces and other surfaces from airborne laser scanning data. For this purpose, different segmentation methods have been developed based, for example, on Hough transform (Maltezos, Ioannidis, 2016), RANSAC (Chen et al., 2012), or surface growing (Alharthy, Bethel, 2004). Kada and Wichmann (Kada, Wichmann, 2012) had proposed point cloud segmentation algorithm called subsurface growing. As indicated by its name, the algorithm is an extension of the wellknown surface growing approach, in which the growing process continues below the surfaces. The purpose was to gather better model data, as roof features and shapes become more apparent regarding these subsurface segments. Despite the fact that the subsurface growing algorithm gives good results, even for complex building roof shapes, the problem remains with roof faces generated by dormers and chimneys. Several dormers and chimneys remain not identified because their shapes are composed of several planar faces. Although the planar segmentation methods of point clouds often serve their purpose, those methods are not suited to segment non-planar objects like power lines, cars, trees and any object with free form shape. There is less works related to non- planar objects segmentation from point clouds. The aim is to combine planar and nonplanar segmentation methods in order to increase entities identification. Clustering is the process of grouping points with similar feature vectors into a single cluster separate from points with dissimilar feature vectors. In the literature, the point clouds clustering methods are usually used for clustering multi-planar segments and 3D building facades (Zolanvari et al., 2018). For this

purpose, large roof segments are extracted with a sub-surface growing method and the remaining points (e.g. dormers and chimneys) with a clustering methods. The Connected Component Analysis, or Labeling (CCA), is a known method of clustering that was widely used in image processing. In this paper, the CCA is exploited for clustering non-planar objects from Airborne Laser Scanning (ALS) LiDAR point cloud. The proposed method does not take only k nearest neighbors point, but rather all points that the distance from each one is less than a fixed distance beforehand. Starting from a random point belonging to initial set of points, this point is labeled with the current label symbol and is removed thereafter from the set. It is different compared from that proposed in the literature (Zhang et al., 2016). Our method does not stop at immediate neighbors of the interest points, but continues until no further neighbors are found. In order to evaluate the clustering algorithm, classification of components is performed. Support Vector Machine, Random Forest and Decision Tree classifiers are tested and compared. The visualization of the results shown the great potential of connected component analysis when applying it for clustering non-planar surfaces such as dormers and chimneys. This paper consists of three sections. Following a description of the methodology, Connected component labeling is introduced. Then, classification step is detailed. Experimental results are visualized followed by discussions. The conclusion summarizes the findings of the present work.

2. METHODOLOGY

This paper focuses on on non-planar surfaces clustering of raw airborne LiDAR data. First of all, sub-surface growing segmentation method is applied on LiDAR point cloud as preliminary filter to extract planar roof surfaces. Then, from untreated points, an adapted algorithm of connected components labeling is applied. Finally, in order to evaluate the clustering algorithm, an interactive resulting components labeling is performed, fol-

^{*} Corresponding author

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 1. Flowchart of the full method

5

7

lowed by classification process. The classifiers Support Vector Machine, Random Forest and Decision Tree are tested and compared. The classification is optimized using grid search and cross validation. The results illustrate the potential of connected component labeling for clustering both of non-planar objects such as roof superstructures and planar surfaces like ground. The flowchart of the proposed method is depicted in Figure 1.

Connected Components Labelling 2.1

Connected components analysis (CCA) ,or connected components labeling, is a clustering algorithm basically used for image processing. In the context of LiDAR data, Milaresis et al. employed CCA to recognize the background and foreground objects from digital elevation model (Miliaresis, Kokkas, 2007). Zhang et al. exploited CCA from 3D point cloud for the classification in urban area (Zhang et al., 2013). in (Xiao et al., 2016), CCA is used for tree points clustering. However, closely 13 return (Components, label(points)) located and intersected trees are clustered together as multi-tree components, which required further treatment with mean shift segmentation. In (Vosselman, 2013), Vosselman exploited knearest neighbors for clustering connected component of point clouds using a fixed neighborhood size.

The proposed CCA does not consider only k nearest neighbors point, but also the neighbors of neighbors. It is different than the one presented in (Zhang et al., 2016), where the authors considered only the immediate neighbors for forming components. The proposed algorithm extends the process until no points are found. The process of CCA is described in algorithm 1.

2.2 Classification

There are two approaches for point cloud classification, pointbased and segment-based classification (Elberink, Vosselman, 2012). In point-based classification approach, class label is affected point by point. For segment-based classification, a prior segmentation method is required before the classification. The class label is assigned by segment and the points of segment share the same class label (Vosselman et al., 2017). In this paper, a clustering of SSG unassigned points is performed using CCA. The resulting components facilitate the interactive labeling step described in paragraph 3.2. For classification, each component is replaced by its points. The considered features are (x, y, z). The Classifiers SVM (Kang, Yang, 2018, Mallet et al., 2011), Random Forests (RF) (Li et al., 2019, Niemeyer Algorithm 1: Connected component analysis

Input: S set of point cloud $\{p_i\} | i = 1..n$ **Input:** *ND* : Neighborhood distance. **Input:** LAB = 0: Label of empty component. **Output:** (Components, *lab*(*points*)) for Each point p_i in S do 1 $lab(p_i) = 0;$ 2 3 while $(\exists p \in S | lab(p) = 0)$ do p : current point to treat; 4 $\hat{list} \leftarrow \{p\};$ $LAB \leftarrow LAB + 1;$ 6 while $(list \neq \emptyset)$ do $c \leftarrow push \ back \ (list);$ 8

```
9
            lab(c) \leftarrow LAB;
            set = {neighbors(c, ND)};
10
            delete c from list:
11
            list \leftarrow list \cup set;
```

et al., 2014) and Decision Tree (DT) (Li et al., 2016, Crasto et al., 2015), widely used for LiDAR data classification are tested and compared.

3. RESULTS AND DISCUSSION

3.1 Study area

The proposed approach is applied on ALS point cloud provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010). It was acquired by Leica Geosystems using a Leica ALS50 system and is characterized by a low density (4-6 points/m²). The area covers urban Vaihingen city in southern Germany. It is marked by the presence of several roof superstructures, as shown in Figure 2.

3.2 Interactive components labeling

For evaluation purposes, the training set is selected, as can be shown in Figure 3. An interactive labeling of components is achieved before the classification. The targeted classes are dormers, chimneys and ground. The components are draped over orthophoto and labeled manually. Than, the label of points is deducted according to the component they belongs to.

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 2. A part of the region of interest in the orthophoto, showing abundance of roof superstructures.



Figure 3. The parts of region forming training and testing sets.

3.3 Data preparation

The learning data set is randomly split in training (70%) and testing (30%) subsets. Classifiers parameters are determined using grid search and cross validation. Table 1 shows the details of each class.

3.4 Clustering results

Figure 4 shows connected component analysis result of unassigned points (Figure 4.(a) using distance of 1.5 m. The resulting components are differently colored to differentiate from each other. The distance is empirically chosen considering the point cloud density (4 to 6 point/m²) and the particularity of the roof superstructures size. From 106,833 unassigned points (Figure 4 (a)), 15,817 components are generated (Figure 4 (b)). In Figure 4.(c), only components with more than five points are considered.

Classes	Designation	Components number	Total points
class 1	dormer	341	5606
class 2	ground	47	1021
class 3	chimney	155	1597



Figure 4. Connected component analysis. (a) Raw points. (b) all resulting components considering neighboring distance of 1.5 m. (c) components with size greater than five points.

Table 1. Detail of each class from selected samples regions

3.5 Evaluations metrics

To evaluate CCA algorithm, the components are subject to classification process. The performance measurements used for classification are overall accuracy (OA), execution time and area under the curve (AUC) score.

Metrics	OA(%)	AUC	Time (s)
SVM	86.22	0.96	11.75
RF	83.93	0.93	2.17
DT	82.26	0.90	0.02

Table 2. Classification of components using different classifiers

3.6 Components classification

The Table 2 illustrates the components classification results. Good scores are achieved by the tested classifiers. The best result is recorded with SVM classifier OA (86.22%) and AUC (0.96). Figure 5 shows multiclass receiver operating characteristic curve for SVM classifier. Classes dormers and ground AUC scores exceeds 0.90 and is close to 0.99 for chimneys class. The figure 6 shows an examples of resulting components projected over the orthophoto. The components represent the classes dormers, ground and chimneys. The neighborhood distance metric directly affects the components results. Testing another distance is time-consuming, due to the components interactive labeling required. This can be seen as drawback of connected component labeling method. The second disadvantage of the CCA method is the dependence of the resulting components to the chosen neighborhood distance. The figure 7 shows two dormers considered as a single due to their proximity. In deed, there is no universal metric usable for all the cases.



Figure 5. Receiver Operating Characteristic Curve for SVM classifier. Class 0: dormers; class 1: ground; class 2: chimneys.

CONCLUSIONS

In this paper connected component labeling for both of planar and non planar objects clustering is proposed. Only a neighborhood distance is initially defined. To evaluate the clustering,









(d)

Figure 6. Component projected over orthophoto. (a) Gable dormers (b)Shed dormers. (c) Ground. (d) Chimneys.

interactive labeling is done, followed by component classification step. Support vector machine, Random forest and decision The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 7. Two dormers considered as a single component due to their proximity.

tree classifiers were exploited. The results demonstrate the relevance of the proposed method to identify dormers, chimneys and ground. Interactive labeling task being long and tedious, only Vaihingen city data set was tested. In future work, further data sets will be studied and compared to the obtained results.

ACKNOWLEDGMENTS

Authors aim to thank the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010) for the furnished Vaihingen city data.

REFERENCES

Alharthy, A., Bethel, J., 2004. Detailed building reconstruction from airborne laser data using a moving surface method. 20th Congress of International Society for Photogrammetry and Remote Sensing, Citeseer, 213–218.

Chen, D., Zhang, L., Li, J., Liu, R., 2012. Urban building roof segmentation from airborne lidar point clouds. *International journal of remote sensing*, 33(20), 6497–6515.

Cramer, M., 2010. The DGPF-test on digital airborne camera evaluation–overview and test design. *Photogrammetrie-Fernerkundung-Geoinformation*, 2010(2), 73–82.

Crasto, N., Hopkinson, C., Forbes, D., Lesack, L., Marsh, P., Spooner, I., Van Der Sanden, J., 2015. A LiDAR-based decision-tree classification of open water surfaces in an Arctic delta. *Remote Sensing of Environment*, 164, 90–102.

Elberink, S. O., Vosselman, G., 2012. Entities and features for classification of airborne laser scanning data in urban area. *IS-PRS Annals Photogramm Remote Sens Spat Information Sciences*, 257–262.

Kada, M., Wichmann, A., 2012. Sub-surface growing and boundary generalization for 3D building reconstruction. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3*, 233–238.

Kang, Z., Yang, J., 2018. A probabilistic graphical model for the classification of mobile LiDAR point clouds. *ISPRS journal* of photogrammetry and remote sensing, 143, 108–123. Li, N., Liu, C., Pfeifer, N., 2019. Improving LiDAR classification accuracy by contextual label smoothing in post-processing. *ISPRS journal of photogrammetry and remote sensing*, 148, 13– 31.

Li, Z., Zhang, L., Tong, X., Du, B., Wang, Y., Zhang, L., Zhang, Z., Liu, H., Mei, J., Xing, X. et al., 2016. A three-step approach for TLS point cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(9), 5412–5424.

Mallet, C., Bretar, F., Roux, M., Soergel, U., Heipke, C., 2011. Relevance assessment of full-waveform lidar data for urban area classification. *ISPRS journal of photogrammetry and remote sensing*, 66(6), S71–S84.

Maltezos, E., Ioannidis, C., 2016. Automatic extraction of building roof planes from airborne lidar data applying an extended 3d randomized hough transform. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 3(3).

Miliaresis, G., Kokkas, N., 2007. Segmentation and objectbased classification for the extraction of the building class from LIDAR DEMs. *Computers & Geosciences*, 33(8), 1076–1087.

Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. *ISPRS journal of photogrammetry and remote sensing*, 87, 152–165.

Vosselman, G., 2013. Point cloud segmentation for urban scene classification. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 1, 257–262.

Vosselman, G., Coenen, M., Rottensteiner, F., 2017. Contextual segment-based classification of airborne laser scanner data. *IS*-*PRS journal of photogrammetry and remote sensing*, 128, 354–371.

Xiao, W., Xu, S., Elberink, S. O., Vosselman, G., 2016. Individual tree crown modeling and change detection from airborne lidar data. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 9(8), 3467–3477.

Zhang, J., Lin, X., Ning, X., 2013. SVM-based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sensing*, 5(8), 3749–3775.

Zhang, Z., Zhang, L., Tong, X., Mathiopoulos, P. T., Guo, B., Huang, X., Wang, Z., Wang, Y., 2016. A multilevel pointcluster-based discriminative feature for ALS point cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6), 3309–3321.

Zolanvari, S. I., Laefer, D. F., Natanzi, A. S., 2018. Threedimensional building façade segmentation and opening area detection from point clouds. *ISPRS journal of photogrammetry and remote sensing*, 143, 134–149.