THE SEGMENTATION OF POINT CLOUDS WITH K-MEANS AND ANN (ARTIFICAL NEURAL NETWORK)

R.A. Kuçak [a], E. Özdemir [a], S. Erol [a,]

[a]ITU, Civil Engineering Faculty, Department of Geomatics Engineering, 34469 Maslak Istanbul, Turkey
kucak15@itu.edu.tr, ozdemiremr@itu.edu.tr, erol@itu.edu.tr

KEY WORDS: Lidar, Laser Scanning, Photogrammetry, Segmentation, K-means, ANN

ABSTRACT:

Segmentation of point clouds is recently used in many Geomatics Engineering applications such as the building extraction in urban areas, Digital Terrain Model (DTM) generation and the road or urban furniture extraction. Segmentation is a process of dividing point clouds according to their special characteristic layers. The present paper discusses K-means and self-organizing map (SOM) which is a type of ANN (Artificial Neural Network) segmentation algorithm which treats the segmentation of point cloud. The point clouds which generate with photogrammetric method and Terrestrial Lidar System (TLS) were segmented according to surface normal, intensity and curvature. Thus, the results were evaluated.

LIDAR (Light Detection and Ranging) and Photogrammetry are commonly used to obtain point clouds in many remote sensing and geodesy applications. By photogrammetric method or LIDAR method, it is possible to obtain point cloud from terrestrial or airborne systems. In this study, the measurements were made with a Leica C10 laser scanner in LIDAR method. In photogrammetric method, the point cloud was obtained from photographs taken from the ground with a 13 MP non-metric camera.

1. INTRODUCTION

Photogrammetric methods are used in geomatics applications for years. Structure from Motion (SfM) (Ullman, S., 1979) became very popular to produce point cloud, which is a photogrammetric approach.

LIDAR, whether from air or ground, is a technical development which enables a large quantity of three dimensional measurements to be collected in a short period of time (English Heritage, 2007).

Segmentation is a process of dividing of the data according to special characteristic features. The most popular algorithm for point cloud segmentation can be considered as k-means (Jain, 2010).

In 2007, Figueiredo and Schnitman used K-means and ANN for image segmentation. They captured photos which are some images of outdoor scenes of sunflower production in different days as well as in different hours of the days. They designed K-means cluster with two clusters that represent the background and foreground regions. On the other hand, they used three clusters for ANN. Their results show that there are a lot of erroneous or undesirable classifications and they obtained lots of discrepancy results. As a result, they said that K-means algorithm has many problems in image segmentation and when K-means and ANN used together for segmentation, good results can be obtained (Figueiredo and Schnitman, 2007).

In this study, we used K-means and ANN segmentation algorithms for the point clouds. Two different point clouds were realized with both techniques and obtained results were compared by using six class for segmentation. These two methods are TLS and Photogrammetry techniques which are widely used for segmentation algorithms. The terrestrial laser scanner system used for this project is Leica C10 and we

preferred 13 MP IPhone 6S Camera, which is a non-metric camera for terrestrial photogrammetry. Also, we used Cyclone Software developed by Leica Geosystem for TLS data and AgiSoft software developed by AgiSoft LLC for Photogrammetry.

2. DATA AND METHODOLOGY

2.1 The Study Area

In this study, ITU (Istanbul Technical University) Yılmaz Akdoruk Student Dormitory's stairs were selected as study area. It is located in Ayazağa Campus of ITU in Turkey. (Figure 1).

Figure 1: Yılmaz Akdoruk Student Dormitory

The study area scanned with Leica C10, which can get 50,000 points per second with 6 mm accuracy until 50 m and measure with impulse method. 3D point cloud of the building processed with Cyclone software by Leica Geosystems. For photogrammetric method, 20 photographs are taken from

approximately 2 meters of distance. AgiSoft by AgiSoft LLC software (trial version) was preferred for photogrammetric point cloud generation through SfM.

2.2 Structure from Motion (SfM)

The fundamentals of Structure from Motion (SfM) (Ullman, S., 1979) are consist of two main fields, Photogrammetry and computer vision. Although SfM is usually used contemporary computer sciences, it is based on Photogrammetry. We can use a number of methods for producing 3D view from two stereo image in photogrammetry. In computer vision, early achievements based on the studies of 3D Scene structure from stereo via an iterative cooperative algorithm (Marr and Poggio, 1976), which searches key points from two images. The methods derived in photogrammetry and earlier computer vision research has an important role for the SfM (Jebara, T. et all, 1999).

Photogrammetric methods have used for years in geomatics applications. Sfm became very popular to produce point cloud, which is a photogrammetric approach. When SfM evaluate as a kind of photogrammetric algorithm, the fundamentals of SfM based on epipolar geometry (Figure 2).
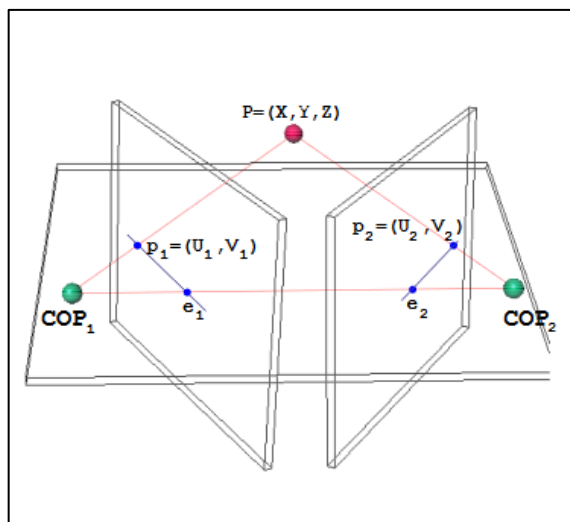

Figure 2: Epipolar Geometry

Figure 2 shows the three points COP1, COP2 and P points in Epipolar plane and the intersection of epipolar line. Assume that, the projections of a point in two image are P1 (u_1 , v_1 , 1) and P2 (u_2, v_2 , 1) , Point P1 in one image, its projection P2 in a second image is restricted to the corresponding epipolar line. The main rule of epipolar geometry means:

$$P_1^T F P_2 = 0 \qquad (1)$$

Equation(1) is typically a least squares minimization and F is fundamental matrix which, constraining to have rank 2 ($\|F\| = 0$), has 9 parameters. Also, this matrix identifies one epipolar line from other image for one point in an image. These perspective projections consist of the fundamentals of SfM algorithms. (Jebara, T. et all, 1999)

2.3 K-means

There are lots of methods for segmentation in literature kd-tree, octree, k-nearest neighbours (KNN) and fixed distance neighbours (FDN) etc. The most popular algorithms for point cloud segmentation are k-means which is known as Lloyd's algorithm (Jain, 2010). K-means minimizes the average of squared distances between the data in the same group. This study was held using MATLAB, which provides k-means++ algorithm. (Vassilvitskii and Arthur, 2007).

The algorithm works as follows;
1) The algorithm chooses preliminary centers using D^2 (D: Distance) weighting.
2) Each point assigned to closest center.
3) Cluster centers calculated again with assigned points via center of mass approach.
4) Steps 2 and 3 are repeated until the centers are stationary.

The data segmented into six with k-means for both TLS and photogrammetry by MATLAB (Figure 2).
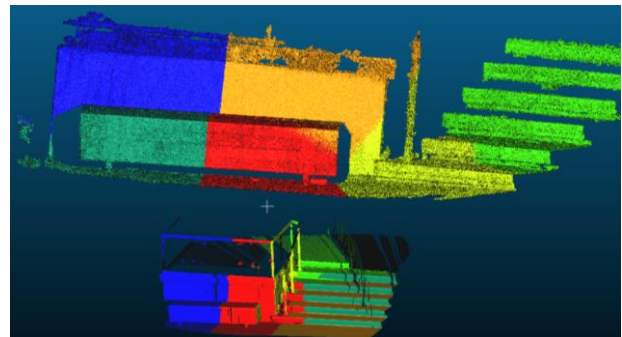

Figure 2: Surface normal, TLS (below) and Photogrammetry (above) K-means Segmentation

The data segmented into six with k-means according to intensity and curvature by MATLAB (Figure 3).
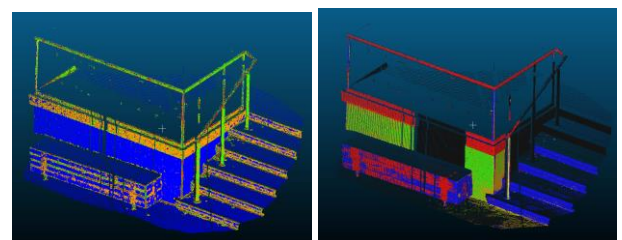

Figure 3: TLS Curvature (Left) and Intensity (Right) k-means Segmentation

2.4 Artificial Neural Networks (ANN)

An ANN is a mathematical method that imitates the functionalities of biological neural networks. The figure 3, shows simple feed forward topology with the input, hidden and output layers (Krenker et al., 2011). In Artifical Neural networks, Patterns are presented to the network via the input layer, which communicates to one or more hidden layers, where the actual processing is done via weighted connections. There are lots of different kinds of learning rules, used by ANN. Feed-forward Artificial Neural Networks, Recurrent Artificial Neural Networks, Hopfield Artificial Neural Network, Elman and Jordan Artificial Neural Networks, Kohonen's Self-Organizing Map etc. (Krenker, Kos, & Bešter, 2011) are some of them. These methods differ according to learning methods. For example; in feed forward back propagation method, information moves only one direction, forward, from the input nodes through the hidden nodes and to the output nodes. There are no cycles or loops in the

network (Auer, Burgsteiner, & Maass, 2008). The ANN used in this study is Kohonen's Self-Organizing Map (SOM), which is widely used for clustering purposes.
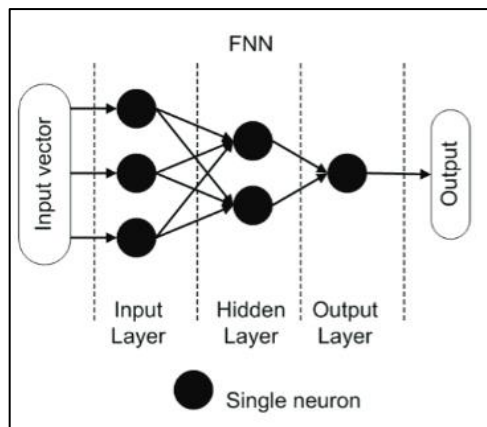

Figure 3: Feed-forward (FNN) artificial neural network (Krenker et al., 2011)

In Figure 3, black lines represent weight layers (w_i). The output of the model (q) with hidden and output layers can be represented by:

$$q = f(\sum w_{j,k} f(\sum w_{i,j} x_i)) \tag{2}$$

where x is the input and f is the transfer function (S. Erol, 2011).

Kohonen's Self-Organizing Map (SOM), is widely used for clustering purposes. The unsupervised training of SOM produces a representation of the input data, called as map. Using this map, the network clusters the input data. The SOM is considered as a software tool that represents high dimensional data in an effective way. This is because it has the ability of transforming complex, high dimensional data into clean geometric associations using low dimensional output. As it compresses the data, it preserves essential topologic and metric connections. The SOM is mostly constructed in a 2D grid of nodes. Each node is related with a model of observations. The models are processed due to analyse the domain of observations. (Kohonen T., 1998)

The SOM's training workflow is represented in the figure below. A neuron's weight is calculated using the equation below.

$$W_v (s + 1) = W_v(s) + \theta(u,v,s).\alpha(s).(D(t)-W_v(s)) \tag{3}$$

Where, W_v is the weights of neuron v, s is the index of the step, $\alpha(s)$ is learning rate, t training sample index, u is the index of the Best Matching Unit (BMU) for D(t), D(t) is the input vector, $\theta(u,v,s)$ is the function of neighbourhood, which calculates distance between neuron u and v for step s. (Kohonen T., 1998).
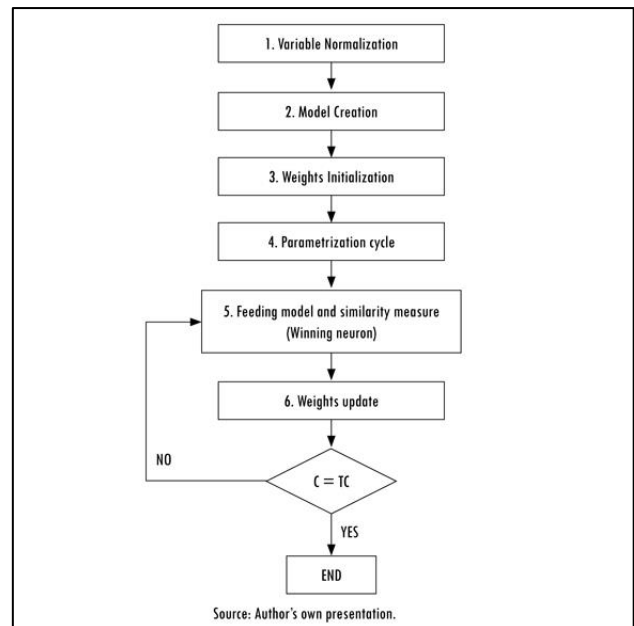

Figure 4: Kohonen's network operation (Kohonen T., 1998)

In this study, a SOM is generated using built-in functions in MATLAB. The network generated six-by-one to produce six clusters (Figure 5). The ANN is trained with the data itself.
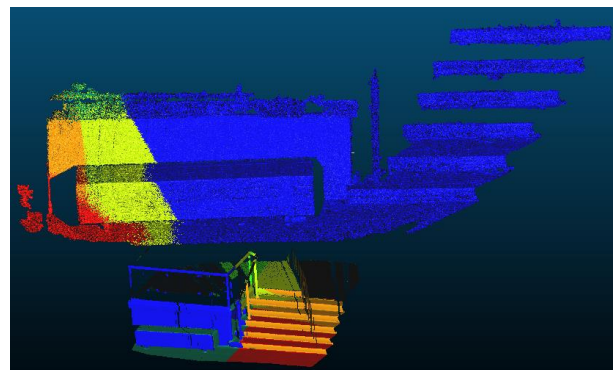

Figure 5: Surface Normal, TLS (below) and Photogrammetry (above) ANN Segmentation

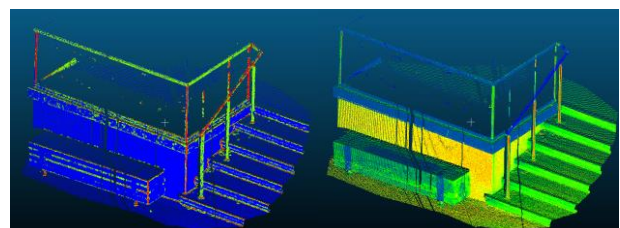The data segmented into six with k-means according to intensity and curvature by MATLAB (Figure 6).

Figure 6: TLS Curvature (Left) and Intensity (Right) ANN Segmentation

3. CONCLUSIONS

In this study, TLS and photogrammetry point clouds are segmented with k-means and ANN. The experiments performed in this study show that each method has advantages as well as disadvantages. According to intensity and curvature, both

methods have good results. We can separate the points according to ground, wall or material types.
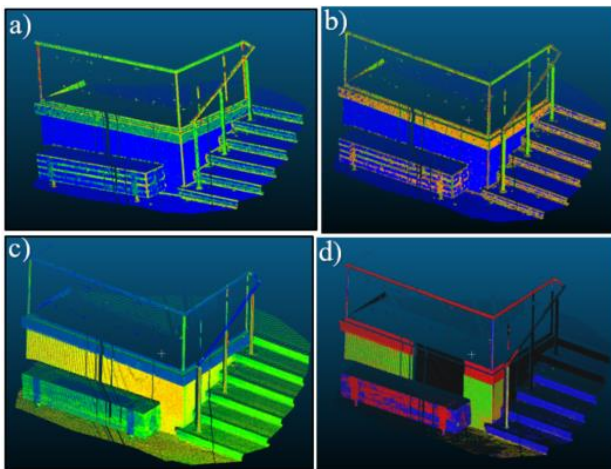

Figure 7: a) ANN curvature b) k-means curvature c) ANN intensity d) k-means intensity segmentation

In Figure 7, curvature segmentation and intensity segmentation have good results because, we can extract corner line with curvature segmentation or extract the surface points belong to wall. Also, ANN results are better than k-means for curvature and intensity segmentation. As seen in Figure 7-d that intensity k-means segmentation has especially minor problem according to ANN intensity segmentation. On the other hand, according to surface normal, we can separate to only corner surfaces and x,y,z direction surfaces. It is not good results for segmentation because we extract wrong information from point clouds (Figure 8). As a result, the main disadvantages observed for k-means can be said that it is not as fast and accurate as ANN. ANN segmentation yields better results than K-means. (Figure 8).
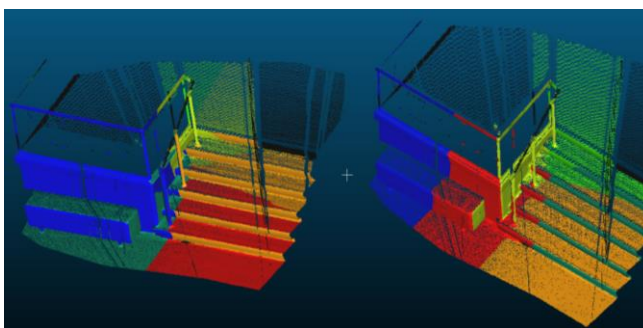

Figure 8: Surface Normal, ANN Segmentation (Left) and K-means Segmentation (Right)

As photogrammetry and TLS methods provide metric data with accuracy information they can be integrated for the segmentation. In this study, photogrammetric data and TLS data provided similar results for same segmentation methods. But, TLS data and photogrammetric data have lots of lack surfaces. So, we didn't extract curvature information from point cloud. If the point clouds are transformed from one to the other and filtered or combine TLS and Photogrammetric data, the segmentation results can be better.

References

Almagro A, Almagro Vidal A., 2007. Traditional Drawings Versus New Representation Techniques. The ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Athens, Greece, Vol. XXXVI-5/C53, pp. 52-57

Arthur, D., & Vassilvitskii, S. (2007, January). K-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms* (pp. 1027-1035). Society for Industrial and Applied Mathematics.

Auer, P., Burgsteiner, H., & Maass, W. (2008). A learning rule for very simple universal approximators consisting of a single layer of perceptrons. *Neural Networks, 21*(5), 786-795.

Domanski, C., 2010. "Properties of the Jarque-Bera Test" FOLIA OECONOMICA 235, Acta Universitatis Lodziensis.

English Heritage, 2007. "3D laser scanning for heritage. Advice and guidance to users on laser scanning in archaeology and architecture" http://archive.cyark.org/temp/EH3dscanning.pdf (02.02.16)

Erol, S. 2011. Time-frequency analyses of tide-gauge sensor data. Sensors, 11(4), 3939-3961.

Figueiredo and Schnitman, 2007. "Using neural network and k-means cluster for image segmentation in outdoor scenes" 2nd International Congress University-Industry Cooperation (UNINDU).

Jain, A. K. 2010. Data clustering 50 years beyond K-means. Pattern recognition letters, 31(8), 651-666

Jebara, T., Azarbayejani, A., Pentland, A., 1999 3D Structure from 2D Motion, MIT Media Laboratory, Cambridge MA,

Marr and Poggio, 1976. Cooperative computation of stereo disparity. Science,194-282

Kohonen, T. 1998. "The self-organizing map" Neurocomputing, 21(1), 1-6.

Krenker, A., Volk, M., Sedlar, U., Bešter, J., & Kos, A. (2009). Bidirectional artificial neural networks for mobile-phone fraud detection. Etri Journal, 31(1), 92-94.

Krenker, A., Kos, A., & Bešter, J. (2011). Introduction to the artificial neural networks: INTECH Open Access Publisher.

Stejakovic V., 2008. "Terrestrial Photogrammetry and Application to Modeling Architectural Objects" Architecture and Civil Engineering Vol. 6, No1, 2008, pp. 113 – 125.

Ullman, S., 1979. "The interpretation of structure from motion" Proceedings of the Royal Society of London B 203, 1979, 405 - 426.