NON-REFERENCE QUALITY EVALUATION FOR INDOOR 3D POINT CLOUDS

Yuhan Lian[1], Chenglu Wen[1], Cheng Wang[1], Jonathan Li[1,2]*

1 Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, Fujian 361005, China
2 Department of Geography and Environmental Management, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
clwen@xmu.edu.cn

ABSTRACT

This paper proposes a novel approach for indoor point clouds quality evaluation, which works well without reference point clouds. In this paper, we mainly evaluate indoor point clouds quality in two aspects: the smoothness of the walls and the degree of occlusion of the walls and the floor. Our approach involves three steps. Firstly, with the S3DIS dataset, we use a deep learning method to train a detector to label walls and floor from indoor scenes. Next, we calculate the normal vector of the wall and the normal vector of each point on the wall. The degree of smoothness of the wall is judged according to the angle between the normal vector of the wall and the normal vectors of the points. Finally, according to the cause of occlusion (objects on the floor or in front of the wall), the occlusion degree of the wall and the floor is obtained. The experimental results demonstrate that the proposed method is suitable for non-reference indoor point cloud quality evaluation.

***Index Terms*—** non-reference quality evaluation, occlusion detection, indoor point clouds

1. INTRODUCTION

With the development of 3D data acquisition sensor, such as stereo cameras, terrestrial laser scanning (TLS), hand-held laser scanning devices, low-cost depth cameras and so on, it has become easier and more affordable to acquire point clouds. Indoor 3D point clouds have essential applications in virtual reality and enhancement, path planning, navigation, building monitoring, etc. However, due to the complexity of the indoor scene and the limited measurement range of the point cloud collection device, there are some quality problems in the acquired indoor 3D point clouds. Many studies have been done in evaluating indoor point clouds quality.

There are mainly two ways to evaluate the quality of point clouds: subjective evaluation and objective measures [1], which are both still open problems. Taking into account the human visual perception of color and shape in 3D color models, Zhang et al. [2] presented a subjective quality evaluation model that can be used to compute the difference between an original 3D color model and the other processed one. Javaheri et al. [3] performed a subjective evaluation of point cloud denoising algorithms and tested the commonly used point cloud objective quality metrics to understand how well they approximate subject evaluations.

In the evaluation of point clouds quality, the objective measurement method should be consistent with the result of subjective evaluation. There are two main ways to evaluate the quality of the image objectively. One is to compare each pixel of the tested image with the pixel of the reference image one by one. The other is to extract the features of the tested image and the reference image, respectively, then compare these features one-to-one. This objective measurement method is a full reference method. However, in the evaluation of point clouds quality, it is difficult to obtain the reference point clouds, because of the characteristics of the sensor itself and the external environment, which make it difficult to objectively evaluate the quality of point clouds.

Huang et al. [4] proposed a method of extracting cloud characteristics in 3D patches, and then using the trained semi-supervised model to evaluate local point cloud quality. Li et al. [5] used different approaches like data noise analysis, geometric characteristics analysis, and registration error analysis to evaluate the quality of indoor point cloud. Alexiou et al. [6] proposed a new objective quality metric based on the angular similarity of associated points belonging to a reference and a point cloud under evaluation.

In this paper, a non-reference quality evaluation metric for indoor point cloud data is presented. Firstly, with the S3DIS dataset [7], we use a deep learning network called pointwise convolutional neural networks [8] to label walls and floor from indoor scenes. Then the indoor point cloud quality is evaluated based on the labeled planes. The evaluation indicators of this method are mainly reflected in two aspects: one is the smoothness of the wall; the other is the occlusion of the walls and the floor. Without reference point clouds, the experimental result is evaluated based on subjective evaluation.

2. METHODOLOGY

The proposed approach contains the following three parts as: (1) Labeling the walls and floor by pointwise convolutional neural networks [8]; (2) Calculating the smoothness of the wall according to the angle between the plane normal vector and the normal vector of the point on the wall; and (3) Evaluating the degree of occlusion of the walls and the floor.

2.1. The walls and floor labeling

In this paper, we mainly evaluate the indoor point clouds quality based on the smoothness of the walls and the degree of occlusion of the walls and the floor, so we first label the walls and the floor. Notice that there already exist some neural networks that have performed well in labeling indoor scenes [8] [9] [10] [11]. Hua et al. [8] presented a convolutional neural network for semantic segmentation and object recognition with 3D point clouds and showed good accuracy on planar structures. In this paper, we use pointwise convolutional neural networks [8] to label the walls and floor.

2.2. The smoothness of the wall

Assuming that a wall is smooth enough, the normal vector of the point on the wall should be parallel to the normal vector of the wall. That is, the angle (θ) between the normal vector of the point, and the normal vector of the wall plane should be 0° or 180°, sin (θ) = 0. Considering the directivity of the vector, we assume that the angle ranges from 0° to 90°. From 90° to 0°, the smaller the angle is, the smaller the value of sin (θ) is, and the smoother the wall is. The angle is closer to 0°; the better point cloud data quality is.

Based on the labeled result, the RANSAC algorithm is used to extract the wall and obtain the normal vector of the wall, $\overrightarrow{\mathrm{n}_0}$. Next, the method proposed by Hoppe et al. [12] as implemented in PCL [13] is applied to compute the normal vector of the point on the wall, $\overrightarrow{n_k}$, which is estimated using the K nearest neighbors of the point. Equation 1 is applied to get the sine of the angle between $\overrightarrow{\mathrm{n}_0}$ and $\overrightarrow{n_k}$, and equation 2 is applied to get the angle. Finally, the smoothness of the wall, S, is computed by the equation 3. In equation 3, N is the number of the point on the wall.

$$\sin < \overrightarrow{\mathrm{n}_0}, \overrightarrow{n_k} > = \sqrt{1 - \left(\frac{\overrightarrow{\mathrm{n}_0} \cdot \overrightarrow{n_k}}{\| \overrightarrow{\mathrm{n}_0} \| \| \overrightarrow{n_k} \|}\right)^2} \tag{1}$$

$$< \overrightarrow{\mathrm{n}_0}, \overrightarrow{n_k} > = \sin^{-1}(\sin < \overrightarrow{\mathrm{n}_0}, \overrightarrow{n_k} >) \tag{2}$$

$$S = \frac{1}{N} \sum_{k=1}^{N} < \overrightarrow{n_0}, \overrightarrow{n_k} > \tag{3}$$

2.3. Occlusion detection

Planar occlusion is caused by the object which locates in front of it, resulting in that the sensor can only collect the information of the object, but cannot collect the planar information. In other words, the point information on the plane is covered by the information of the occlusion. If the points of the occlusion are projected straightly to the plane, a more complete plane can be restored. Based on this feature, the occlusion of the plane can be estimated. The method to detect the occlusion of indoor scenes involves five steps. Without losing the generality, we take the floor plane as an example.

Firstly, based on the labeled result, a RANSAC algorithm is used to extract the floor. Then, given a minimum boundary rectangle from (x_{min}, x_{max}) and (y_{min}, y_{max}), the block size b_z is set and then a two-dimensional coordinate system is built. A two-dimensional coordinate system is defined as:

$$\mathrm{grid}[u, v]: \begin{cases} 0 \leq u \leq \left\lceil \dfrac{x_{max} - x_{min}}{b_z} \right\rceil \\ 0 \leq v \leq \left\lceil \dfrac{y_{max} - y_{min}}{b_z} \right\rceil \end{cases} \tag{4}$$

Next, equation 5 is applied to divide points (x, y, z) on the floor into corresponding grids. Let N_1 be the number of occupied grids. Similarly, the number of occupied grids is obtained in the same way for all the points in the indoor scene, including the points of the occlusion. Let N_2 be the number of occupied grids for all points. Finally, equation 6 is applied to compute the degree of the occlusion, ρ.

$$\begin{cases} u = \left\lfloor \dfrac{(x - x_{min})}{b_z} \right\rfloor \\ v = \left\lfloor \dfrac{(y - y_{min})}{b_z} \right\rfloor \end{cases} \quad (5)$$

$$\rho = \frac{N_2 - N_1}{N_2} \quad (6)$$

3. RESULTS AND DISCUSSIONS

The experimental data is from S3DIS dataset [7], which is collected from six large-scale indoor areas that originate from three different buildings of mainly educational and office use. In our experiment, we take three different indoor scenes as examples. These three scenes (a hallway, a conference room, and a lounge) with different degrees of occlusion are shown in Fig. 1. Since there is no reference data, we evaluate the objective measurement results based on subjective evaluation.

(a) hallway

(b) conference room

(c) lounge

Fig. 1. Three illustrations of indoor scenes [7]

Figure 2, 3, and 4 show the planes extracted from the three indoor scenes separately.

Fig. 2. The extracted planes of the hallway

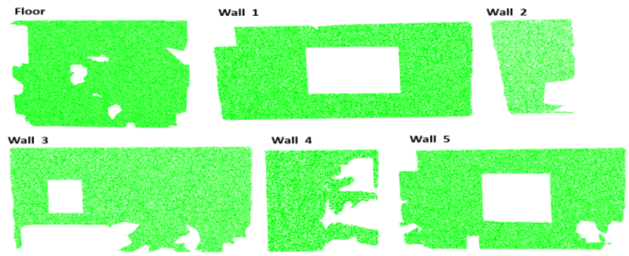

Fig. 3. The extracted planes of the conference room

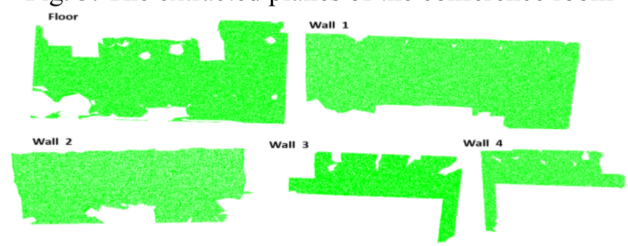

Fig. 4. The extracted planes of the lounge

As can be seen from Table 1, with the increase of K, the angle between the normal vector of the point on the plane and the normal vector of the plane will become smaller and smaller. It shows that the smaller K value can reflect the local smoothness. It is in line with our subjective evaluation. In the subjective evaluation, wall 3 is smoother than wall 2 in Fig. 5, which is consistent with the objective measurement of Table 1.

Scenes	Typical planes	Smoothness S (°, degree)			
		K=10	K=15	K=20	K=30
hallway	Wall 1	2.84	2.75	2.70	2.62
	Wall 2	16.60	7.21	5.72	3.52
	Wall 3	2.18	1.65	1.21	1.15
	Wall 4	2.57	2.36	2.21	1.93
conference room	Wall 1	3.03	2.15	1.82	1.66
	Wall 2	4.47	3.18	2.86	2.43
	Wall 3	1.86	1.73	1.63	1.47
	Wall 4	5.52	4.78	4.52	4.12
	Wall 5	1.98	1.84	1.75	1.62
lounge	Wall 1	2.09	1.93	1.82	1.67
	Wall 2	6.88	6.09	5.77	5.55
	Wall 3	9.84	7.54	6.56	5.99
	Wall 4	6.73	6.50	6.34	6.08

Table 1: Different K nearest neighbors of the point on the plane

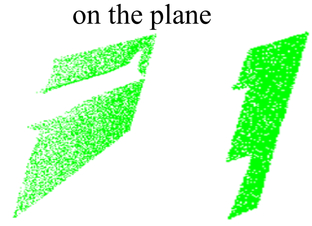

Fig. 5. Local information of the walls: a is wall 2 of the hallway, b is wall 3 of the hallway

In the subjective evaluation, the hallway has the least degree of occlusion, and the lounge has the largest degree of occlusion in the occlusion of the floor. Table 2 demonstrates that feature numerically. Fig. 6 shows some results of occlusion detection of the floor.

Scenes	Typical planes	Occlusion percentage		
		b_z= 0.02 m	b_z= 0.05 m	b_z= 0.08 m
hallway	Floor	0.0600	0.0309	0.0271
	Wall 1	0.1001	0.0935	0.0857
	Wall 2	0.1692	0.1862	0.1761
	Wall 3	0.0763	0.0765	0.0801
	Wall 4	0.1537	0.1457	0.1412
conference room	Floor	0.1203	0.0876	0.0742
	Wall 1	0.1929	0.1813	0.1578
	Wall 2	0.1625	0.1515	0.1462
	Wall 3	0.1897	0.1734	0.1555
	Wall 4	0.2183	0.1855	0.1973
	Wall 5	0.2068	0.1880	0.1714
lounge	Floor	0.2946	0.2647	0.2502
	Wall 1	0.1830	0.1593	0.1605
	Wall 2	0.1754	0.1508	0.1244
	Wall 3	0.2083	0.2010	0.2246
	Wall 4	0.1294	0.1386	0.1676

Table 2: The different block size b_z for occlusion evaluation

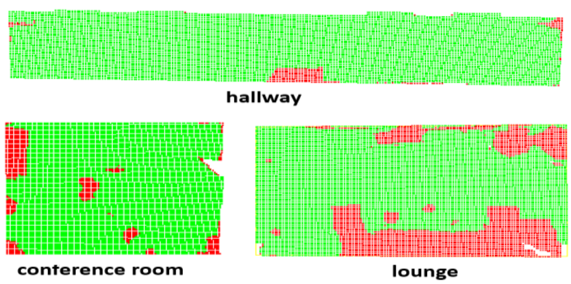

Fig. 6. Occlusion detection of the floor, the red areas represents the occlusion

4. CONCLUSIONS

In this paper, we evaluate non-reference indoor point clouds quality in terms of both smoothness and occlusion. Our experiments show that the proposed method can be used as a criterion for evaluating indoor point clouds quality without reference point clouds.

REFERENCES

[1] Dumic E, Duarte C R, da Silva Cruz L A. Subjective evaluation and objective measures for point clouds—State of the art[C]//Smart Grid Metrology (SmaGriMet), 2018 First International Colloquium on. IEEE, 2018: 1-5.

[2] Zhang J, Huang W, Zhu X, et al. A subjective quality evaluation for 3D point cloud models[C]//Audio, Language and Image Processing (ICALIP), 2014 International Conference on. IEEE, 2014: 827-831.

[3] Javaheri A, Brites C, Pereira F, et al. Subjective and objective quality evaluation of 3D point cloud denoising algorithms[C]//Multimedia & Expo Workshops (ICMEW), 2017 IEEE International Conference on. IEEE, 2017: 1-6.

[4] Huang F, Wen C, Luo H, et al. Local quality assessment of point clouds for indoor mobile mapping[J]. Neurocomputing, 2016, 196: 59-69.

[5] Li G, Chen Y, Wen C, et al. Quality evaluation of point cloud model for interior structure of a common building[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2017 IEEE International. IEEE, 2017: 6040-6043.

[6] Alexiou E, Ebrahimi T. Point cloud quality assessment metric based on angular similarity[C]//International Conference on Multimedia and Expo (ICME). 2018 (CONF).

[7] Armeni I, Sener O, Zamir A R, et al. 3d semantic parsing of large-scale indoor spaces[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1534-1543.

[8] Hua B S, Tran M K, Yeung S K. Pointwise convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 984-993.

[9] Qi C R, Su H, Mo K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[J]. Proc. Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, 1(2): 4.

[10] Maturana D, Scherer S. Voxnet: A 3d convolutional neural network for real-time object recognition[C]//Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on. IEEE, 2015: 922-928.

[11] McCormac J, Handa A, Davison A, et al. Semanticfusion: Dense 3d semantic mapping with convolutional neural networks[C]//2017 IEEE International Conference on Robotics and automation (ICRA). IEEE, 2017: 4628-4635.

[12] Hoppe H, DeRose T, Duchamp T, et al. Surface reconstruction from unorganized points[M]. ACM, 1992.

[13] Rusu R B, Cousins S. 3d is here: Point cloud library (pcl)[C]//Robotics and automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 1-4.