[1]Zhiheng Ding
[1]Renfu Li

Point Cloud Reduction Method Based on Curvature Grading and Voxel Filtering

Journal of Electrical Systems

Abstract: *-* In the context of automated assembly engineering based on reverse engineering in the aerospace field, redundant data will make it impossible to complete 3D model reconstruction efficiently and quickly, and point cloud data must be streamlined. A single simplification method cannot accurately retain the feature information of the scanned point cloud. Aiming at this background, a point cloud simplification method combining curvature grading and octree voxel filtering is introduced. The obtained curvature is divided into different levels by using the properties of logarithmic function, and the strong and weak feature points are distinguished. The weak feature points are filtered by octree voxel, and the filtered results are merged with the strong feature points to obtain the final streamlined results. This method is compared with the random sampling method and the point cloud simplification method combining 3D-SIFT feature extraction and voxel filtering. The experimental results show that this method not only greatly reduces the amount of data, but also preserves the local details of the original point cloud data, so as to achieve the purpose of efficiently compressing point cloud data.

Keywords:Reverse engineering; point cloud reduction; curvature grading; octree; informationentropy evaluation

I. Introduction

In the field of spacecraft assembly, reverse engineering is often used for virtual assembly. Nowadays, laser scanning is widely used to quickly collect 3D coordinates of an object's surface [1]. Under the condition of sufficient computing power, not removing redundant points will result in the waste of computing resources. However, in engineering applications with limited computing power, redundant data will have a bad effect on the speed of data processing. Therefore, it is necessary to propose a reasonable and efficient method to extract feature points from point cloud data. At present, the methods of point cloud reduction are mainly divided into two parts: mesh reduction and direct reduction, among which random sampling, uniform sampling and non-uniform sampling are the most popular classical methods [2]. The principle of random sampling is very simple [3]. Firstly, the number of sampling points is determined, and the random removal points are sampled. The uniform sampling principle is similar to voxelized grid sampling method, which has the advantage of high accuracy, but its time complexity is increased. Non-uniform sampling uses geometric features to remove redundant points pertinently, and randomly chooses different reduction strategies to reduce the number of point clouds. However, the above classical methods mainly focus on the integrity and efficiency of the method, so it will lead to a large number of feature point information loss.

In recent years, domestic and foreign scholars have carried out a lot of research on reduction methods to solve the problem of data redundancy. Zhao fuqun et al adopt the point cloud simplification method of point importance judgment, which keeps the details effectively, but the evaluation index is complex and the computation complexity is high [4]. Li Haipeng et al adopt the method of FPFH feature extraction. This method preserves the feature information of the model as much as possible. It has high efficiency and strong applicability, but it has too many preset parameters and is not good for small-scale point cloud simplification [5]. Qin Jianguo et al used the multi-scale feature aggregation module to adaptively aggregate the local and global features of point cloud blocks of different scales, and used LSTM to aggregate the features of point cloud blocks of different scales, so as to better retain the local features, estimate more accurate local detail normals [6]. However, this method takes a long time to calculate, and LSTM aggregate scale block feature is more inclined to local detail feature, so it needs to consider how to keep the optimal solution on the accuracy of local detail feature and global feature. Liu et al use the octree filtering method based on K-means clustering, which is too much computation, only retain a part of the feature points effectively, and has great limitations [7].

In order to overcome the high computational complexity and many preset parameters in the reduction method, based on the traditional point cloud reduction method, a new method is proposed to distinguish the feature regions of different grades based on curvature classification, the point cloud can be simplified by using different strategies for different regions. This method can extract the feature information of point cloud data keenly and retain the feature points to a great extent.

[1]School of Aerospace Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

II. POINT CLOUD REDUCTION METHOD BASED ON CURVATURE CLASSIFICATION AND OCTREE METHOD

A. Curvature Value Estimation

The curvature of point cloud data is calculated by quadric surface fitting. In the process of surface fitting, it is necessary to estimate the normal vector at the point to be found, and then set up the local space coordinate system at the point, and its depth direction is the estimated normal vector at the point. In this stage, the first step is used to get the space points in the local coordinate system, and the curvature of the points is calculated from the fitted quadric surface parameters. Principal component analysis (PCA) is usually used to estimate the local normal vector of point cloud data [8][9][10]. For any point P_i of the input point cloud, there will be a microtangent plane at that point, which is obtained by fitting k-neighborhood at the point P_i. The linear approximation equation of the microtangent plane is as follow.

$$ax + by + cz - d = 0 \quad (1)$$

where a, b and c are the normal vectors of the microtangent plane, which are consistent with the estimated normal vector at point P_i. The distance between the tangent plane and the origin of the coordinate system is denoted by d. The three-dimensional coordinates of the midpoint of the microtangent plane are x, y and z, respectively.

The matrix M is covariance matrix, calculated by using the k-nearest neighbors at thepoint P_i. It can be expressed as follows.

$$M = \begin{pmatrix} \sum_{j=1}^{k} \Delta x_j \Delta x_j & \sum_{j=1}^{k} \Delta x_j \Delta y_j & \sum_{j=1}^{k} \Delta x_j \Delta z_j \\ \sum_{j=1}^{k} \Delta y_j \Delta x_j & \sum_{j=1}^{k} \Delta y_j \Delta y_j & \sum_{j=1}^{k} \Delta y_j \Delta z_j \\ \sum_{j=1}^{k} \Delta z_j \Delta x_j & \sum_{j=1}^{k} \Delta z_j \Delta y_j & \sum_{j=1}^{k} \Delta z_j \Delta z_j \end{pmatrix} \quad (2)$$

where the average values of the coordinates of the points in the neighborhood are expressed as $\bar{x}, \bar{y}$and$\bar{z}$. $\Delta x_j = x_j - \bar{x}$, $\Delta y_j = y_j - \bar{y}$, $\Delta z_j = z_j - \bar{z}$. The letter $j$denotes the number of the neighborhood interior point, j=1, 2, 3 … k. The estimated value of the normal vector (a, b, c) at the point P_i is the eigenvector corresponding to the minimum eigenvalue of the matrix M.

The direction of the normal vector estimated by Principal Component Analysis is not consistent, so it is necessary to judge the angle between adjacent normal vectors. If the angle is less than $\pi/2$, it is reversed to ensure that all normal vectors point in the same direction [11].

The quadratic surface S (u, v) near the fitting point $P_i$in the local coordinate system is obtained by the following formula (3) .

$$\begin{cases} S(u,v) = \lfloor u, v, w(u,v) \rfloor \\ w(u,v) = \theta_1 u^2 + \theta_2 uv + \theta_3 v^2 + \theta_4 u + \theta_5 v \end{cases} \quad (3)$$

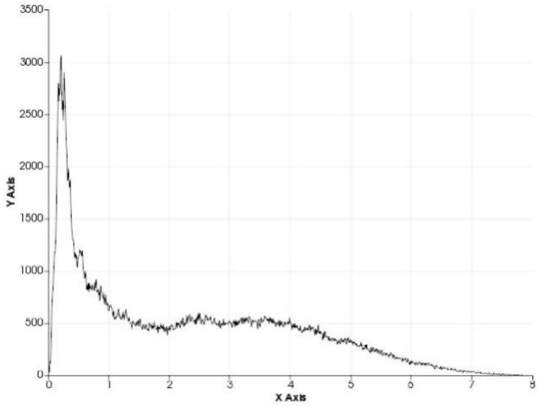
where θ_1, θ_2, θ_3, $\theta_4$and θ_5 are the best parameters of quadric surface. After solving the parametric equation, the average curvature $\bar{\rho}_i$at the point P_i is calculated according to the curvature property of the parametric surface as follows.

$$\overline{\rho_i} = \theta_1 + \theta_3 \quad (4)$$

B. Curvature distribution characteristics of point clouds

Curvaturevalueisanimportantgeometricparameterinformationinthesurfacecharacteristicsofpointclouddata. Thesizeofthecurvaturevalueatacertainpointcandirectlyreflectthedegreeofsurfacechangenearthepoint. Inthispaper, theStanfordChinadragonpointcloudisselectedastheexperimentalobject, whichhasalargescaleandobvioussurfacecharacteristics. AsshowninFigure1, a)andb)aretheoriginalpointcloudvisualizationanditscorrespondingpoints-curvaturevaluedistributionmap, respectively. ThehorizontalaxisXrepresentsthesizeofthecurvaturevalue, andtheverticalaxisYrepresentsthecorrespondingnumberofpointclouds. AscanbeseenfromFigure(1), thepointcloudsunderdifferentdatacharacteristicsmainlycontainsmallcurvatureareas, andthelargecurvatureareaswithobvioussurfacechangesaccountforasmallproportion. Therefore, whensimplifyingthedata, thecurvaturelevelcanbedividedaccordingtodifferentlevelstoselectoutthelargelycompressedsmallcurvatureareasand preservetheoriginaldetails.

(a) China Dragon visualization



(b) point number-curvature distribution of China Dragon

Figure 1Chinadragon curvature characteristics

C. Curvature Classification Strategy

According to the geometric properties of the logarithmic curve, the curvature classification strategy of this paper is: according to the natural logarithmic function to calculate the curvature of a point of the characteristic grade. The greater the curvature, the higher the corresponding grade [12]. It is necessary to normalize the curvature values of different point cloud data to a fixed range because the curvature distribution of different data sets is different. According to the geometric properties of the logarithmic function, the rate of change of the value of the function will decrease with the increase of the independent variable. At the same time, the natural logarithm value of the change rate can not be too small, in order to make curvature classification more differentiated, and can make the grade of curvature changes with sensitivity. Therefore, we will use the following steps to normalize the calculated curvature values in section 1. 1 to the range of 0~8. The detailed process is as follows.

(1) Traversing all the data points, the curvature of the point cloud data is calculated. ρmaxrepresents the maximum curvature. $\rho_{\min}$ represents the maximum curvature.

(2) The normalized curvature value $H_i$of each point is calculated, whose value is normalized to the range of 0~8. The formula for the normalization step is as follows.

$$H_i = \frac{8 \cdot \left(\overline{\rho_i} - \rho_{\min}\right)}{\rho_{\max} - \rho_{\min}} \quad (5)$$

The normalized curvature values need to be graded. The numerical formula for the classification is as follows.

$$D_i = ceiling\left\lfloor 2\ln\left(\frac{H_i+1}{H_0+1}\right)\right\rfloor$$
$$D_i = \begin{cases} 0, & \text{if} D_i \le 0 \\ D_i, & \text{if} 0 < D_i < 5 \\ 5, & \text{if} D_i \ge 5 \end{cases} \quad (6)$$

where *ceiling*[] represents an up-rounded function. D_i is the curvature grade of the point P_i. H_0 is the given curvature threshold. According to Formula (6), the curvature value is divided into six grades from 0 to 5. The region in which the points are normalized to a level of 0 is recorded as a very flat region. The curvature of the points in this region is less than the threshold H_0. At the same time, the region whose D_i is less than or equal to 3 is considered as non-critical feature region, while the region whose D_i is greater than or equal to 3 is considered as critical feature region. The selection principle of threshold H_0 is that there is a local extremely flat region in different point cloud data sets, and the threshold H_0 is determined by calculating the average curvature of all points in the region. When the compression ratio is small, H_0 can be taken as 0 directly.

The interval of the curvature value between the first and second order is denoted as ΔH_1. In the same way, the interval between the third and fourth grades is denoted as ΔH_2. According to the investigation and analysis of a large number of data, the curvature of more points in the point cloud data is small curvature value, but the proportion of the points in the area of large curvature value, i. e. sharp and inflection points, is less. Therefore, the properties of logarithmic function can effectively reflect the relationship between the number of large curvature and small curvature in point clouds.

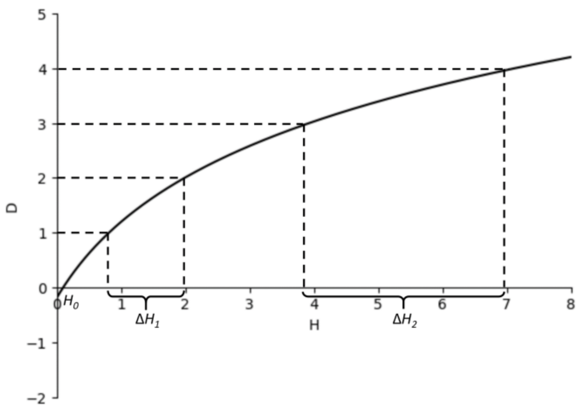


Figure 2 curvature classification sketch based on logarithmic function

D. Octree Voxel Filtering

Octree structure is a data model. It divides the geometric objects in a 2n×2n×2n 3D by a recursive loop to form a direction graph with root nodes [13]. For a space object of size 2n×2n×2n, it needs to be partitioned n times at most. The result is a directional graph in which each partition contains a root node.

The improved octree method used in this paper is a different method to deal with point cloud data. The idea of traditional octree voxel filtering is to use voxel centroid instead of all points in the grid. However, doing so will result in distortion of the sampled data because the sampled points are not the actual data points in the original point cloud. To solve this problem, the improved octree method uses the point closest to the center of gravity to replace all the points inside the voxel.

E. Streamline the Method Flow

First, the local normal vector of point cloud data is estimated by principal component analysis, then the local surface is fitted by the point of K-nearest neighbor and the estimated normal vector of the point at a certain point, and the mean curvature at that point is calculated. Classification of curvature values based on logarithmic function. The points with large curvature are extracted and saved as a new set of point clouds. The part of the characteristic point cloud is eliminated in the original point cloud, and the remaining non-characteristic point cloud is preserved. Then the uncharacteristic region is filtered by using the improved octree filter. Finally, the key point cloud with high curvature and the filtered point cloud with low curvature are combined and spliced by

PCL library, and the new point cloud is the result of simplification. Figure 3 shows a flowchart of the overall streamlined approach.

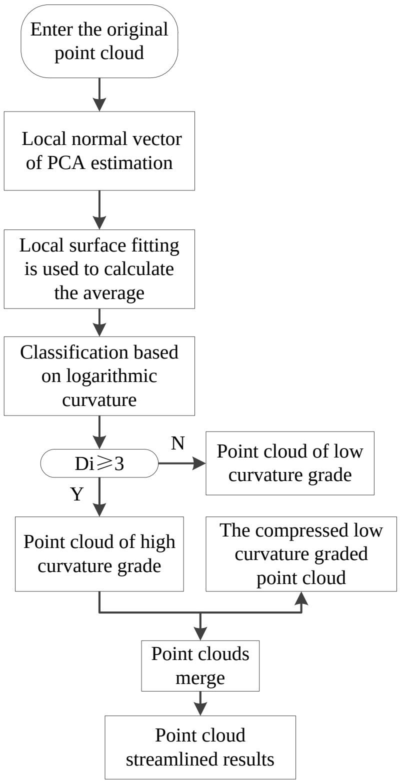
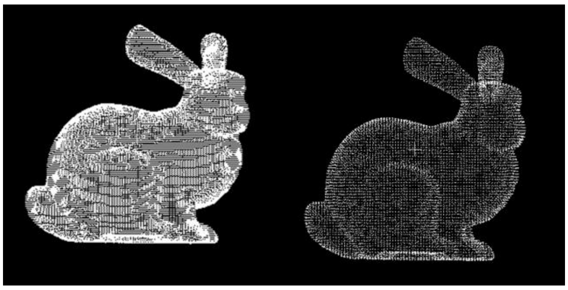


Figure 3 A simplified method flow chart based on curvature classification combined with voxel filtering

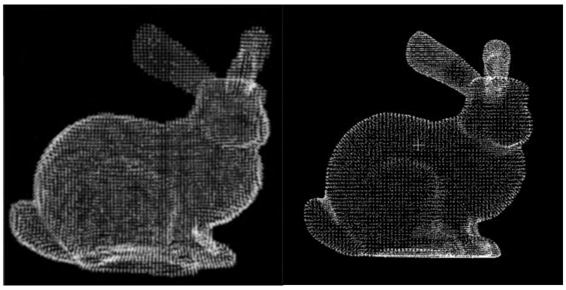
III. Experiment

In this paper, we select the Stanford Bunny Point Cloud with complicated surface and obvious feature change. The number of points in the Bunny Point Cloud is 35947. This data set contains abundant information of surface features, can effectively test the compression effect of this method.

By visualizing the result of Point Cloud reduction, the ability of different methods to extract detail features is compared and analyzed. Figure 4 shows a simplified visualization of the original point cloud with three different compression methods at the same compression rate. In addition to the method in this paper, the other two methods are point cloud compression using random sampling method and point cloud compression using 3D-SIFT feature extraction method in literature [14] . The figures (b), (c) and (d) in figure 4 below show the simplified visualizations of three different compression methods, each with a compression of 50% . According to the visualization results, the points obtained by the random sampling compression method are uniformly distributed in each region, and the difference between the regions of characteristic points and non-characteristic points can not be visualized. The method of feature extraction based on 3D-SIFT in literature [14] and the method proposed in this paper can obviously show that the points in the feature areas such as ears and feet of bunny point cloud are dense, and the feature information can be well preserved.



a)Original Point Cloud b)Random Sampling Something



c)Literature [14] Algorithm d)This Article Method

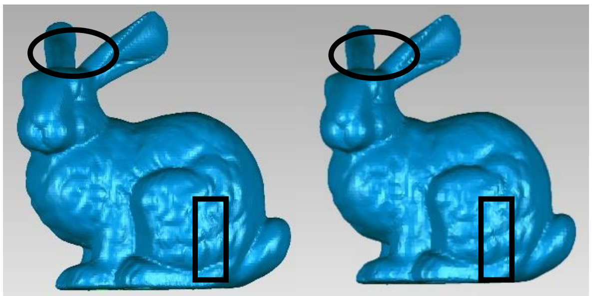
Figure 4Bunny Point Cloud visualizes the results in different ways

IV. Analysis of Experimental Results

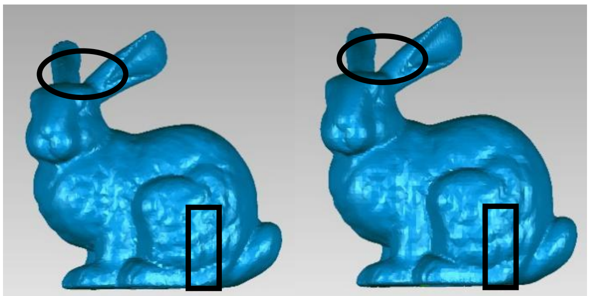
At home and abroad, scholars mainly use visual effects to evaluate the accuracy of point cloud data compression, and some scholars use information entropy to quantify the simplified results. In this experiment, we visualized the simplified results of point cloud and calculated the information entropy of Point Cloud to evaluate and Quantitative analysis the compressed results.

A. Visual effects assessment

Using the compression results obtained by the three methods in Figure 4 above, triangulation was performed in Geomagic Studio 2014 software, and the results were visualized and compared with the original point cloud triangulation results, the aim of this paper is to verify the superiority of the proposed compression method in the surface reconstruction step. The result is shown in Figure 5.



a) The original point cloud triangle is partitionedb）Randomly sampled triangular faces are partitioned



c)Triangulation of Reference [14] methodd)Triangulation of this paper method

Figure 5 Three-dimensional visualization models with different methods

It can be concluded from the content shown in figure 5 that the part of the Bunny Point Cloud delimited in the circle is the characteristic region of the point cloud, which has distinct edges and angles and complex surface changes. In the three-dimensional model built by the method presented in this paper, the difference between the part of circle and the original data is the smallest. In the model of random sampling and facet slicing by reference [14] method, the effect of the area reconstruction is rough, the feature is not well preserved, and the serrated texture appears in the circle. It shows that these two methods are not accurate enough for the feature extraction of this region. The subcharacteristic regions of the Bunny point cloud are marked by rectangular boxes. It can be seen that the smoothness of the proposed method is closest to the original point cloud, followed by random sampling. The reconstruction results obtained by the method in literature [14] are the worst. For other flat regions, such as the back of the Bunny Point Cloud, the proposed method will filter the data points of this region in a high proportion when it is compressed, so that the triangular patches formed in this region are larger when the model is reconstructed. However, the surface of this area is relatively smooth, and only a few points can accurately restore the surface shape of this area.

B. Information Entropy Analysis

By using curvature to solve information entropy, its value can objectively and quantitatively evaluate the quality of the simplified results. The higher the entropy value, the more information there is, and the more disordered the point cloud distribution is, the more obvious the characteristics are [15]. Therefore, the information entropy of point cloud data is used in this paper to objectively and quantitatively evaluate the quality selection of three different reduction methods. The formula for entropy at a given point is as follows.

$$C_i = -p_i \log_2 p_i - \sum_{j=1}^{k} p_j \log_2 p_j \quad (7)$$

$$p_i = \frac{\rho_i}{\rho_i + \sum_{j=1}^{k} \rho_j} \quad (8)$$

$$p_j = \frac{\rho_j}{\rho_i + \sum_{j=1}^{k} \rho_j} \quad (9)$$

where $i$represents the number of the entire point cloud point. i=1, 2, 3⋯N. The information entropy of the point i is represented by C_i. $\rho_i$is the curvature at the point i. The point numbers in the neighborhood of the point $i$are denoted by j. j=1, 2, 3⋯k. $\rho_j$is the curvature at the point j. The probability distribution of curvature of point i is expressed by p_i. The probability distribution of curvature of point j is expressed by p_j. The total entropy of each point in the point cloud is calculated as the entropy of the whole point cloud. The whole point cloud entropy is calculated by the following formula.

$$C = \sum_{i=1}^{n} C_i \quad (10)$$

In order to evaluate the effect of point cloud data simplification under different compression ratios, the entropy values of the three methods are calculated respectively, as shown in Table 1.

Table 1 The information entropy values of the results of different reduction methods under different compression ratios

	Random sampling	Literature [14] method	This article method
Compression ratio 50%	58735. 8	58760. 4	58801. 8
Compression ratio 60%	47402. 8	46901. 2	46727. 4
Compression ratio 70%	34937. 1	35043. 4	35653. 4

The sampling points of the three methods are not exactly the same under the same compression ratio although there is little difference. In order to eliminate the influence of the small number of points on the experimental results, this paper uses the average information entropy to carry out a more accurate quantitative evaluation of the reduced results [14] .

$$\overline{C} = \frac{C}{N} \quad (11)$$

where N represents the total number of points contained in the resulting point cloud for each method. Table 2 below shows the average information entropy values of the three reduced methods at different compression ratios. The larger the average entropy value is, the more information is contained in the unit volume, the better the effect of feature preservation is.

Table 2 The average information entropy values of the results of different reduction methods under different compression ratios

	Number of point clouds	Random sampling	Number of point clouds	Literature [14] method	Number of point clouds	This article method
Compression ratio 50%	18020	3. 25948	18005	3. 26356	18007	3. 26550
Compression ratio 60%	14583	3. 25055	14400	3. 25703	14329	3. 26104
Compression ratio 70%	10803	3. 23401	10800	3. 24476	10930	3. 26200

It can be seen from Table 2 that the average information entropy of the proposed method is higher than that of the other two methods at the same compression ratio. With the increase of compression ratio, the average information entropy values of the three methods show a decreasing trend at different compression ratios. According to the data in the Table 2, the average entropy of the random sampling method decreases most significantly with the increase of the compression ratio. The average information entropy of the proposed method is stable at each compression ratio, and it does not decrease significantly at high compression ratio.

V. CONCLUSION

Aiming at the problem of redundant data in point cloud data, a method of curvature classification based on logarithm function and octree voxel filtering is proposed. The simplified results obtained by this method are compared with other methods by point cloud visualization analysis and 3D modeling visualization analysis, and the performance of the method is analyzed subjectively. The curvature-based mean entropy is introduced to evaluate the ability of the Quantitative analysis method to retain the feature points objectively. According to the experimental results, the average information entropy is only 0. 107% different when the compression ratio is 70% and the compression ratio is 50%, to facilitate subsequent surface reconstruction operations.

REFERENCES

[1] Liang Yan. Research on 3D laser scanning point cloud data reduction and modeling application [D]. China University of Mining and Technology, 2021.

[2] Li Peipei, Cui Fengying. Optimization of curvature-graded point cloud data reduction based on dichotomous K-means clustering [J]. Electronic Measurement Technology, 2022, 45(04):66-71.

[3] Guo Jiafu, Shi Lei, Yan Limin. A method of 3D point cloud data reduction based on strong feature point reservation [J]. Industrial Control Computer | Ind ContrlComput, 2020, 33(11):92-93.

[4] Zhao Fuqun, Li Yinglei. Point Cloud simplification based on point importance judgment [J]. Computer Systems & Applications, 2023, 32(09):197-202.

[5] Li Haipeng, Xu Dan, Fu Yuting, et al. A method of scattered point cloud reduction based on FPFH feature extraction[J]. Journal of Graphics, 2022, 43(04):599-607.

[6] Qin Jianguo, Hu Xiaoyang. Research on point cloud de-noising method based on PCPNet[J]. Industrial Control Computer | Ind ContrlComput, 2023, 36(04):107-108+111.

[7] Liu M J, Zhao J R, Guo X F， et al. Application of Improved Point Cloud Streamlining Algorithm in Point Cloud Registration[C]// 2020 Chinese Control And Decision Conference (CCDC), April 22-24, 2020, Hefei, China.

[8] Fu Shifeng, Jia Xiaoliang, An Lei, et al. Prediction method of aero-engine overhaul cycle based on PCA-BP[J]. Aeronautical Computing Technique, 2022, 52(03):62-66.

[9] Chen Yi, Wang Yong, Li Jinlong, et al. Efficient Point Cloud registration algorithm based on principal component analysis[J]. Laser Optoelectronics Progress, 2023, 60(14):376-383.

[10] Ma Weifeng, Wang Jinliang, Zhang Jianpeng, et al. Point Cloud feature extraction based on improved normal vector estimation[J]. Science of Surveying and Mapping, 2021, 46(11):84-90+146.

[11] Chen Changwen, Pan Guobing, Zhang Shuai, et al. Tunnel axis extraction based on PCA normal vector estimation and RANSAC algorithm[J]. Coal Geology of China, 2023, 35(02):61-66.

[12] Li Jintao, Cheng Xiaojun, Yang Zexin, et al. Point cloud data compression method based on curvature classification[J]. Laser & Optoelectronics Progress, 2019, 56(14):248-255.

[13] He L P, Yan Z M, Hu Q J, et al. Rapid assessment of slope deformation in 3D point cloud considering feature-based simplification and deformed area extraction[J]. Measurement Science and Technology, 2023, 34(5).

[14] Xing Ying, Song Tao, Zhao Yan, et al. 3D-SIFT feature extraction combined with voxel filtering for point cloud reduction[J]. Laser Journal, 2023, 44(03):163-169.

[15] Zhu Guangtang, Ye Minlv. Point Cloud denoising and quantitative evaluation based on curvature feature[J]. Bulletin of Surveying and Mapping, 2019(06):105-108.