A New Method for Cloud Data Reduction Using Uniform Grids

Zhao Sanyuan, Li Fengxia, Liu Yongmei, Rao Yonghui

School of Comp. Sci. & Tech, Beijing Institute of Tech, Beijing 100081, China
sanyuan.zhao@gmail.com

Abstract - Due to the continuous development of 3D scanning technology, the point cloud data is becoming more and more redundant. Storing or handling these data will consume a lot of time and computer resources. In this paper, a method for cloud data reduction based on 3D Grids is proposed. It includes searching k-nearest neighbors for constructing data topology, calculating tangent plane normal, sampling points using uniform 3D grids and selecting points that satisfy the requirements. The method retains more features than traditional methods, and it avoid the estimating of curvature. Experiments by different point cloud data using the proposed method and other methods show that the proposed method can reserve more minutiae while remove the redundant data.

Index Terms - Reverse engineering, octree, angel between normal, k-nearest neighbors.

I. Introduction

In reverse engineering, 3D scanner as a main tool is widely used, it can get the 3D point cloud data without touching the object, thus completing the reconstruction of the model. Point cloud can also be called unorganized data set, there is no geometric topology relationship between the points, and it's a simple three-dimensional set of points. These points are defined by the x, y and z coordinates. The point cloud data is usually dense and scattered.

The purpose of simplifying the point cloud is retaining more model information with fewer points. The methods of simplifying point cloud are classified into two categories in accordance with whether to build a mesh structure of point cloud: mesh-based methods and point-based methods.

Mesh-based methods have to build the triangular mesh of the points, and then reduce the number of triangles according to application-dependant criteria. It is usually complex and time-consuming.

Point-based methods avoid building triangular mesh of the point cloud. They are more simple and faster. These methods include random sampling method, bounding box method, uniform grid method and so on. Random sampling method is the simplest and easiest method. It generates an integer that smaller than the total number of points each time, and then deletes the point corresponding to the integer. This method has a large randomness and a relatively poor effect.

The bounding box[1][2] method starts by building a grid structure. The input points are assigned to the corresponding cell. For each cell that is not empty, the point which is nearest to the centre of the cell will be selected to represent all the points in this cell. Martin et al. [3] proposed a data reduction method using uniform grid. This method uses a "median filtering" approach, which has been widely used in image processing. Not the same with the bounding box method, in uniform grid method, a median point is selected to represent points belong to the cell that is not empty.

Bounding box method and uniform grid method overcome the limitation of the random sampling method, but there is a drawback due to the use of uniform size grids that can be insensitive in capturing a part shape.

The curvature method[4] is another point-based simplification method, this method find out k-nearest neighbors for each point firstly, and then fit a surface using these points. The curvature of the surface can be seen as the estimated value of the point. The area where the curvature is relatively great should retain more points. This method reserves more details in feature areas, but the curvature calculation[5] needs complex matrix operations, it is time-consuming especially when dealing with large point cloud data.
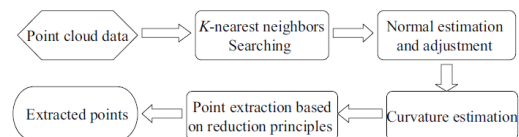

Fig.1 The procedure of the curvature method

Except the previous methods, K. H. Lee et al. [6] presented a non-uniform grid reduction method that the size of the grids can be varied based on part shape. This method is more effective to freeform surfaces than the uniform grid method, but still difficult to identify the boundary points. Fengxia Li et al. [7] proposed a method that adopts nonlinear least square for point-cloud simplification.

Traditional methods, such as random sampling, uniform grids and bounding box do not take into account the local characteristics of the point cloud data, therefore cannot retain the details of the raw point cloud. The curvature method can preserve the minutiae using the estimation of the curvature, but the curvature calculation needs to consume too much time. In this paper, a new point cloud data reduction method that can overcome the limitations of the previous methods is proposed.

II. A new reduction method using uniform 3D grids

The new method includes six steps:
1. Searching the K-nearest neighbors for each point.
2. Unit normal vector calculation for each point.

3. V' calculation, is introduced in the following text.
4. Sampling points using 3D grids.
5. Sampling points using threshold of
6. Output the sample points.

A). Basis of the method

As we all know, for the points from a surface, the normal directions are almost same in flat area, but very different in bending area. Therefore, the angle between normal directions could be used in the reduction method.

Cosine of the angle between two unit vectors can be denoted by their dot product. In this paper, the mean value of dot products of the point's normal vector and its neighbor's normal vector is selected as an element to determine whether the point is sampled. This mean value is called V in this paper.

Let the unit normal vector of point Pi be (vx_i, vy_i, vz_i) , Its K-nearest neighbors' unit normal vector are, $(NX_1, NY_1, NZ_1), \cdots ,$ (NX_k, NY_k, NZ_k), then the V of Pi is:

$$V = (\sum_{i=1}^{k} | nx_i * NX_i + ny_i * NY_i + nz_i * NZ_i |) / k \qquad (1)$$

Assuming that the point cloud data in local area is smooth, the normal vector angel between Pi and its neighbor is not greater than 90 degrees, the calculated dot product is non-negative. So we use the absolute value of the dot product for calculation.

The calculated V values between 0 and 1, the smaller V means the greater degree of bending in the local area of that point. Use $V' = 1 - V$ instead of V for subsequent calculations. The smaller V' means the local area is more flat.
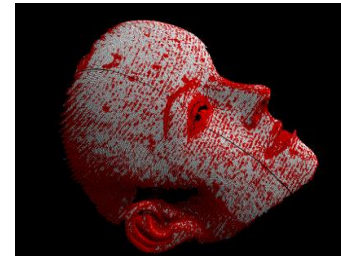
In the Fig.2, the points are marked in different colors. The points whose V' is greater than the specified threshold are marked in red, the others are marked in white. It can be seen that red points are located in bending areas mostly.
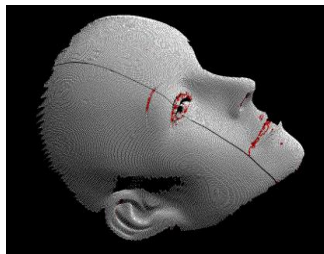
a(0.1)
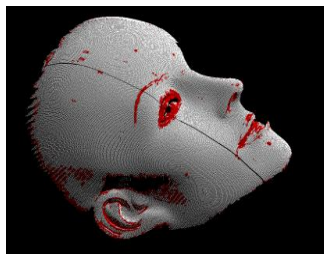
b(0.01)

c(0.001)

Fig.2 The comparison using different thresholds of V' to distinguish between points. The figure in brackets is the value of threshold

B) . Searching the K-nearest neighbors

Using octree structure can significantly reduce the time of searching the k-nearest neighbors for each point.

Octree is defined as: if the tree is not empty, the number of children of any node in the octree is zero or eight. Octree is a hierarchical tree structure formed based on recursive decomposition. It's the expansion of the quadtree structure. The principle of the octree structure is same as binary tree and quadtree. Octree is initially used in the field of 3D solid modeling [8], thereafter it's widely used as an efficient index in three-dimension space.

The root node is firstly created when create an octree. The node represents the smallest cube that encloses all points of the point cloud. Then the cube is divided into eight smaller sub-cubes, each corresponds to a child node of the root. As shown in Fig.3, split the child cubes by this rule until there is no need to split again. Every leaf node of the octree represents a smallest division of the space. Each point of the point cloud should be inserted into the corresponding leaf node.
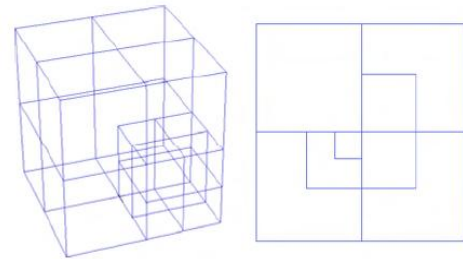

Fig. 3 The octree structure

When searching for the K-nearest neighbors of Pi, only points in the same leaf node or the adjacent leaf nodes are need to be searched.

C). Unit normal vector calculation

Point cloud data obtained by 3D scanner usually containing the normal vector and the reflection intensity information. The point cloud data obtained by optical scanner, in addition to the spatial coordinate values, doesn't contain other information generally. The unit normal vector of point should be calculated in the program.

Since the normal vector accuracy requirement is not so

high in our data reduction method, we use the principle component analysis (PCA) that proposed by Hoppe et al. [9] in 1992.

Assume that the sample surface is smooth everywhere, therefore the local K-neighbors of any point can be well fitted with a plane. For any point Pi, after obtaining its k-neighbors by octree structure, we can estimate a tangent plane Tp(Pi). The plane Tp(Pi) can be expressed as below:

$$Tp(Pi) = \text{argmin}_{(n,d)} \sum_{i=1}^{k} (n \cdot P' - d)^2 \qquad (2)$$

In Eq.2, n is the normal vector of plane Tp(Pi), d is the distance of Tp(Pi) to the origin of the coordinates, P' is the centroid of the K-nearest neighbors.

Let the tangent plane Tp(Pi) pass through the point P', the normal vector n satisfies |n|=1, our goal can be transformed into the eigenvalue decomposition of the positive semi-define covariance matrix M in Eq.3.

$$M = \sum_{i=1}^{k} (Pi - P')(Pi - P')^T / K \qquad (3)$$

If λ_i denote the absolutely smallest eigenvalue of M associated with eigenvector V_i, then normal vector could be either V_i or $-V_i$. It's usually needed to adjust the normal direction of every point. But the calculation of V in Eq.1 is build on the assumption of the angel between normal is not greater than 90 degrees, it doesn't need to know the true direction of normal. We use the absolute value of the dot product of normal vectors to avoid the adjustment of the direction.

After estimating the unit normal vector, we can calculate the value of V using Eq.1 and so V' can be known.

D). Sampling points using 3D grids

If sample the points using V' directly, the point has a V' greater than the threshold should be sampled. The threshold is user-defined. The points in flat areas will be removed on a large scale. As shown in Fig.2, the white points will be removed and so generate holes.

So, the sampling process is divided into two steps. The first step is sampling points using uniform 3D grids.
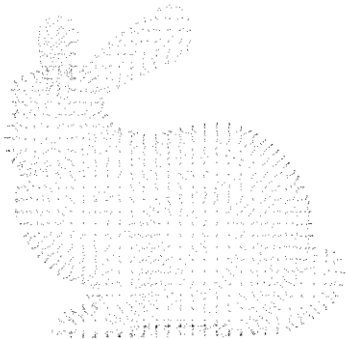

Fig. 4 sampling points using uniform 3D grids

Use the bunny point cloud as an example, the original point cloud has 35947 points. Fig.4 is the result after sampling points using 3D grids. The total number of cells after creating bounding box is 262144. For each cell that not empty, use the 'median filtering' approach to calculate a median point for sampling. There are 2619 points are sampled, it means that 259525 leaf nodes are empty.

After the sampling, the resulting points distribute uniformly, but lots of detail information is lost.

E). Sampling points using threshold of V'

The second step of sampling process is sampling points using threshold of V'. For each point that not be sampled in step 1, compare its V' with the threshold. The value of threshold is user defined. If the point's V' is greater than the threshold, it should be sampled. The user can control the reduction rate by changing the threshold of V'.

When the threshold is 0.03, the total number of points after sampling is 6215. The value of threshold is between 0 and 1. No point is selected for sampling if the threshold of V' is greater than 1 while all points are sampled if the threshold is smaller than 0.
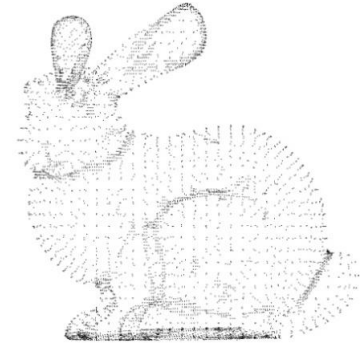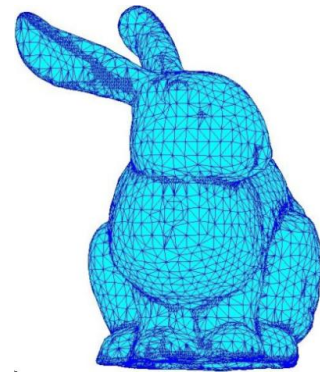

Fig. 5 sampling points using threshold of V'

Fig. 6 triangular mesh model of the streamlined point cloud

Fig.5 is the result after the second sampling. The points sampled this time mostly distribute in the intersection of different surfaces. These areas are very important in the three-dimensional reconstruction of point cloud. Fig.6 is the triangular mesh model using the streamlined bunny point cloud.

After this step, the reduction operation is finished, output the streamlined point cloud in the end.

III. Experimental results

The point cloud of car model is shown in Fig.7(a), in different surfaces, the density changes a lot. Fig. 7(b, c, d, e) also shows the reduction result of car model using different methods. In non-feature regions, our method sample points uniformly and avoid the generation of holes. In feature areas, our method sample more points to retain the minutiae of the surfaces.
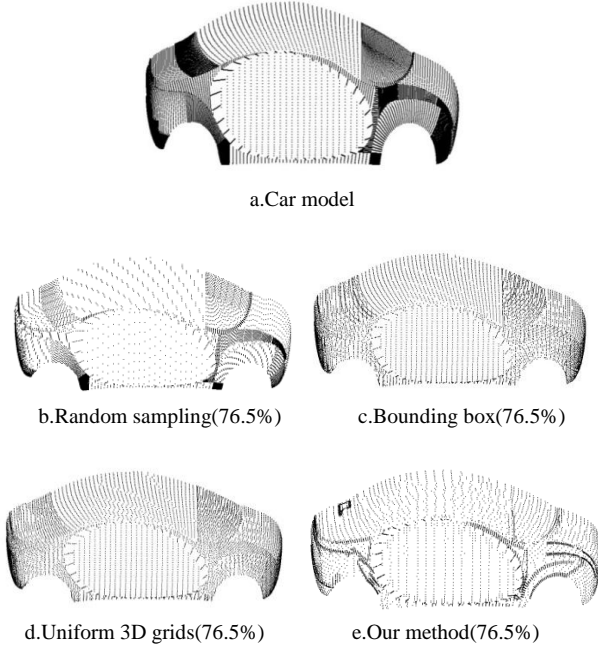

a.Car model

b.Random sampling(76.5%)   c.Bounding box(76.5%)

d.Uniform 3D grids(76.5%)   e.Our method(76.5%)

Fig. 10 Reduction result of the car model using different methods, figures in   brackets are reduction rate.

IV. Conclusions

A point cloud data reduction method based on uniform grids is proposed in this paper, which reduces the amount of point cloud data by using the part geometry characteristic. The method includes two sampling process. The first sampling uses 3D uniform grids to get the general outline of the point cloud. The second sampling uses the angel between normal to increase the model's details in feature regions. The methods absorb the advantages of both uniform grids and curvature method, and avoid the shortcomings of curvature method in surface fitting. The method can also achieve different reduction rate by specifying threshold. The experiments show that this method can achieve good results.

References

[1] Sun W, Bradley C, Zhang Y F et al., Cloud Data Modeling Employing a Unified Non-redundant Triangular Mesh Computer-Aided Design, 2001, 33(2):183-193.

[2] D. Filip, R. Magedson, R. Marko, Surface algorithms using bounds on derivatives. Computer Aided Geometric Design, 1986, 3(2):295-311.

[3] R.R. Martin, I.A. Stroud and A.D. Marshal. Data Reduction for Reverse Engineering. RECCAD, January 1996(1068):63-69.

[4] X. Du and Y. Zhuo. A point cloud data reduction method based on curvature. In IEEE 10th International Conference on Computer-Aided Industrial Design, 2009, pp. 914–918.

[5] T. Surazhsky, E. Magid, O. Soldea, G. Elber, and E. Rivlin. A comparison of Gaussian and mean curvatures Estimation methods on triangular meshes.Taipei: proceedings of 2003 IEEE International Conference on Robotics & Automation, 2003, pp. 1021-1026

[6] Lee K H, Woo H, Suk T, Point Data Reduction Using 3D Grids. The International Journal of Advanced Manufacturing Technology, 2001, 18(3): 201—210.

[7] Fengxia Li, Rong Tang, Chen Liu and Haikun Yu. A Method for Object Reconstruction Based on Point-cloud Data via 3D Scanning, 2010 International Conference on Audio Language and Image Processing (ICALIP), 2010, pp. 302 – 306.

[8] D. MEAGHER, Geometric Modeling Using Octree Encoding. Computer Graphics and Image Processing. 1982, 19(2): 129-147.

[9] H. Hoppe, T. DeRose, T. Duchamp, et al., Surface reconstruction from