Journal of Southeast University (English Edition) Vol.20 No.1 Mar. 2004 ISSN 1003—7985

Segmentation of scattered point data through a new curvature analysis algorithm

He Meifang Zhou Laishui Zhang Liyan Liu Shenglan

(Research Center of CAD/CAM Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: A systematic scheme is proposed to automatically extract geometric surface features from a point cloud composed of a set of unorganized three-dimensional coordinate points by data segmentation. The key technology is a new algorithm that estimates the local surface curvature properties of scattered point data based on local base surface parameterization. Then, eight surface types from the signs of the Gaussian and mean curvatures provide an initial segmentation, which will be refined by an iterative region growing method. Experimental results show the scheme's performance on two point clouds.

Key words: point data; segmentation; local base surface parameterization; eight surface types; curvature

Segmentation has been an essential part in the process of surface modeling from scanned point data[1]. It is the process of portioning a point cloud into meaningful regions or extracting important features from the point data. The majority of point data segmentation methods can be classified into two categories: edge-based methods, face-based methods.

The edge-based methods attempt to detect discontinuities in the surfaces that form the closed boundaries of components in the point data. Ref.[2] used a semi-automatic edge-based approach for orthogonal cross-section (OCS) models. Surface differential properties were estimated at each point in the model, and the curvature extremes were identified as possible edge points in the model. Then an energy-minimizing active contour was used to link the edge points. Ref.[3] used a parametric quadric surface approximation for segmentation of 3-D measured data. They used a least-square method to fit those parametric quadric surfaces. The emphasis of this paper is that the local curvature of the measured points is more accurately reflected by the fit quadric surface than by other methods[2,4] (Dauborx frame). Once an accurate determination of local curvature is made, the edges can be determined[5].

All of these methods above are based on two assumptions. One requires that the points be structured in a special pattern[3,6], the other supposes that the relationship between discrete points is described by an other model[2,4,5]. Few data segmentations have been directly to dense scattered points. The main impedance to automation is the overlapping and disorganized nature of scattered points.

In this paper, data segmentation directly employs the unorganized point cloud. We improve on the local parametric-quadric-surface-approximating method of Ref.[3] by using local base parameterization to extend the segmentation technology directly to scattered point data. Based on a new curvature analysis algorithm, the proposed segmentation scheme is composed of four successive processes: spatial partition, building the K-neighbour at each point, first identification of flat points, and the second identification of the other seven surface points. The main contributions of this paper are:

1) Based on an LBS parameterization method, an algorithm of curvatures of randomly measured points is proposed.

2) Data segmentation technique is extended to dense scattered point data by using the new curvature algorithm based on the developed LBS parameterization method.

3) Two extractions improve the efficiency of coarse segmentation. Points that belong to planes are firstly extracted using the distance between a point and its tangent plane. The remainders are secondly extracted by a complex calculation using the developed curvature algorithm.

1 Definition

In reverse engineering, data points are collected through a contact or noncontact digitizer. A data point is normally represented by its three coordinates x, y and z. Let x_i be the i-th measured data point, and $X =$

Received 2003-09-27.

Foundation items: Aeronautical Science Foundation of China (No. 00H52069) and the Natural Science Foundation of Jiangsu Province (No.BK2001408).

Biographies: He Meifang (1972—), female, graduate; Zhou Laishui (corresponding author),male,doctor,professor, zlsme@nuaa.edu.cn.

$\{x_1, x_2, \cdots, x_N\}$ be the whole set of data points. We first introduce some preliminary definitions, which are useful for subsequent discussions.

1) Parametric quadric surface (PQS) All the possible parameterized surfaces of the first and second degree in 3-D space are called as PQS. A specific surface in PQS is denoted as S. The equation of PQS is

$$\boldsymbol{r}(u, v) = \sum_{i=0}^{2}\sum_{j=0}^{2}\boldsymbol{Q}_{ij}u^i v^j \tag{1}$$

or the matrix form is

$$\boldsymbol{r}(u, v) = \{1, u, u^2\}\,\boldsymbol{Q}\begin{Bmatrix}1\\ v\\ v^2\end{Bmatrix} \tag{2}$$

where $\boldsymbol{Q}$ is a 3×3 matrix with vector-valued elements $\boldsymbol{Q}_{ij}$. Components of $\boldsymbol{r}$ and $\boldsymbol{Q}$ are

$$\boldsymbol{r}(u, v) = \{x(u, v), y(u, v), z(u, v)\}$$

$$\boldsymbol{Q} = \{\boldsymbol{Q}_{ij}\} = \{a_{ij}, b_{ij}, c_{ij}\}$$

The vectors

$$\boldsymbol{W} = \{u^0v^0, u^0v^1, u^0v^2, u^1v^0, u^1v^1, u^1v^2, u^2v^0, u^2v^1, u^2v^2\}^{\mathrm{T}}$$

$$\boldsymbol{a} = \{a_{00}, a_{01}, a_{02}, a_{10}, a_{11}, a_{12}, a_{20}, a_{21}, a_{22}\}^{\mathrm{T}}$$

in $\mathbf{R}^9$ are introduced and the vectors $\boldsymbol{b}$, $\boldsymbol{c}$ are defined similarly. Then, the components of Eq. (1) can be written as

$$\boldsymbol{x} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{a},\ \boldsymbol{y} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{b},\ \boldsymbol{z} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{c}$$

2) K-neighbour There are k points in X which are the nearest to a measured point x_i. These k points which include x_i are called as K- neighbour, noted as Nbhd(x_i).

3) LBS Short for a local base surface. It is the surface that approximates the underlying surface, and that should satisfy the condition that any two points on the underlying surface should have two different projected points on it.

2 New LBS Parameterization Method

If the measured points are randomly spaced, it becomes difficult to locate values of u, v and to use the local parametric-quadric-surface-approximating method to calculate the underlying surface curvature. Based on the affine ratio theory[7], the LBS parameterization method is proposed. Firstly, K-neighbour of measured points are projected onto LBS. The parameters of the projected points in LBS are then used as the parameters of its K-neighbour in the local parametric quadric surface. The parameterization procedure includes the following steps:

1) LBS creation A local base surface can usually be defined from some characteristic curves or a plane approximating the underlying geometry[8]. It can be of any form, either truly freeform, or a particular simpler form, such as a planar, cylindrical or spherical surface. However, it should satisfy the condition that any two points on the underlying surface should have two different projected points on it. Here, the tangent plane (TP) which approximates K-neighbour at a measured point x_i is regarded as the LBS of its K-neighbour.

2) Projection The projection may be done either normally to the LBS, or according to a given projection. In this paper, the projection is done normally to the LBS.

3) Directions of u, v The direction of u is obtained by drawing the line L_1 from the projected point x_i' of a measured point x_i to the projected point p' which is the furthest to x_i' in the LBS. The direction of v is got by drawing the line L_2 from the point x_i' which is perpendicular to the line L_1 (see Fig.1 (a)).

4) A parameter set $\{(u_j, v_j), j=1, 2, \cdots, k+1\}$ Linking x_i' and each projected point p_j' of K-neighbour, and k lines are got. The set of k lines is $\{L(x_i', p_j'), j=2, 3, \cdots, k+1\}$. Then, projecting $L(x_i', p_j')$ $(j=2, 3, \cdots, k+1)$ to u, the projecting distance set $\{d_j, j=2, 3, \cdots, k+1\}$ is got. By sorting these distances (see Fig. 1(b)), the maximum and minimum of these distances are obtained, $d_{\max}$ is the value of the maximum distance, and $d_{\min}$ is the value of the minimum. The parameter u set $\{u_j, j=1, 2, \cdots, k+1\}$ of K-neighbour is got by

$$u_j = \frac{d_j - d_{\min}}{d_{\max} - d_{\min}} \qquad 2\leqslant j\leqslant k+1,\ u_1 = 0.0 \tag{3}$$

The parameter v set $\{v_j, j=1, 2, \cdots, k+1\}$ of K-neighbour is got similarly. The parameter set $\{(u_j, v_j), j=1, 2, \cdots, k+1\}$ is got by matching two parameter sets of u, v.

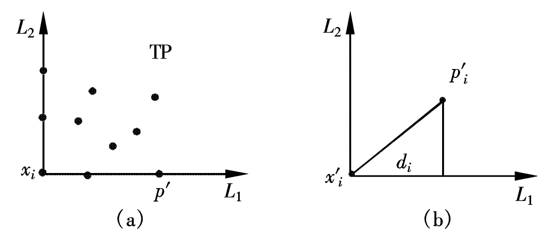


Fig.1 LBS parameterization. (a) The directions of L_1 and L_2; (b) The distance of d_i

3 Curvature Analysis Algorithm for Scattered Points

In order to use a quadric surface approximating K-neighbour at each point to estimate the curvature of each point, a parameterization is realized by projecting the measured points (K- neighbour) normally to an

LBS. The parameters of the projected points are then used as the parameters of the measured points (K-neighbour). Further, the least-square approximate scheme minimizes the sum of the actual Euclidean distance between the neighborhood data points and the parametric quadric surface. The Gaussian and mean curvatures are computed from the locally approximated surfaces.

1) Parameterization Using the above LBS parameterization method, u, v values of K-neighbour of each point are obtained.

2) Calculating the PQS which approximates Nbhd (x_i) The least-square scheme minimizes the sum of the squares of the actual Euclidean distances between Nbhd (x_i) and the corresponding points on the parametric quadric surface. According to Ref. [5], $\boldsymbol{Q}$ of Eq.(3) can be given as

$$\boldsymbol{Q} = (\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M})^{-1}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{Z} \tag{4}$$

where

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{W}_0^{\mathrm{T}} \\ \boldsymbol{W}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{W}_K^{\mathrm{T}} \end{bmatrix}$$

$$\boldsymbol{Z} = [\boldsymbol{Z}_x, \boldsymbol{Z}_y, \boldsymbol{Z}_z] = \begin{bmatrix} x_0 & y_0 & z_0 \\ x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_k & y_k & z_k \end{bmatrix}$$

where $\boldsymbol{M}$ is a $(k+1)\times 9$ matrix, $k > 8$ is required for local surface approximation; $\boldsymbol{Z}$ is the matrix of x_i and Nbhd (x_i). Because $\boldsymbol{M}$ and $\boldsymbol{Z}$ are known, $\boldsymbol{Q}$ can be calculated. For calculating the invert of $\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}$, this paper refers to Ref. [9]. By operating Eq.(4), $\boldsymbol{Q}$ is obtained.

3) The Gaussian (K) and mean (H) curvatures of the i-th point x_i The matrix $\boldsymbol{Q}$ is substituted into Eq.(2) to obtain the local quadric surface $\boldsymbol{r}(u, v)$. The partial derivatives $\frac{\partial \boldsymbol{r}}{\partial u}$, $\frac{\partial \boldsymbol{r}}{\partial v}$, $\frac{\partial \boldsymbol{r}}{\partial u \partial u}$, $\frac{\partial \boldsymbol{r}}{\partial u \partial v}$ and $\frac{\partial \boldsymbol{r}}{\partial v \partial v}$ are noted as $\boldsymbol{r}_u$, $\boldsymbol{r}_v$, $\boldsymbol{r}_{uu}$, $\boldsymbol{r}_{uv}$, and $\boldsymbol{r}_{vv}$. The unit normal vector $\boldsymbol{N}$ is then given as

$$\boldsymbol{N} = \frac{\boldsymbol{r}_u \times \boldsymbol{r}_v}{\|\boldsymbol{r}_u \times \boldsymbol{r}_v\|}$$

and the Gaussian curvature K and the mean curvature H can be given as

$$K = \frac{LN - M^2}{EG - F^2} \tag{5}$$

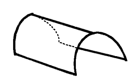
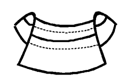
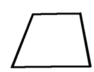
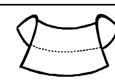
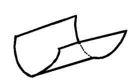
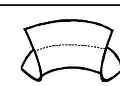
$$H = \frac{EN + GL - 2FM}{2(EG - F^2)} \tag{6}$$

where $E = \boldsymbol{r}_u \cdot \boldsymbol{r}_u$, $F = \boldsymbol{r}_u \cdot \boldsymbol{r}_v$, $G = \boldsymbol{r}_v \cdot \boldsymbol{r}_v$, $L = \boldsymbol{r}_{uu} \cdot \boldsymbol{N}$, $M = \boldsymbol{r}_{uv} \cdot \boldsymbol{N}$, $N = \boldsymbol{r}_{vv} \cdot \boldsymbol{N}$.

4 Eight Fundamental Surface Types

Arbitrarily smooth surface can be subdivided into simpler regions of constant surface curvature sign based on the signs of the mean and Gaussian curvatures at each point[10]. There are eight possible surface types surrounding any point on a smooth surface based on surface curvature sign: peak, pit, ridge, valley, saddle ridge, saddle valley, planar (flat), and minimal. These simple surface types are shown in Tab.1.

Tab.1 Eight fundamental surface types from surface curvature sign

H	$K>0$	$K=0$	$K<0$
$H<0$	 Peak	 Ridge	 Saddle ridge
$H=0$	None	 Flat	 minimal
$H>0$	 Pit	 Valley	 Saddle valley

5 Coarse Segmentation Scheme

Based on the new curvature algorithm and eight fundamental surface types, which are labeled by the signs of mean and Gaussian curvatures above, a coarse segmentation scheme is developed.

5.1 Spatial partition

Spatial partition strategy is employed in order to improve the efficiency of this approach. First, a data file is read into the memory of a computer and 3-D coordinate points are saved into the 1-D array noted as X; at the same time, we set the maximums and the minimums of x, y and z. According to these maximums and minimums a cube that includes all points of the point cloud and parallels coordinate axis is developed. Second, based on the quantity and the dense of the point cloud, the cube is divided into $m \times n \times l$ cubes. Then, each point is judged as to which cube includes it. Its serial number is added in the list corresponding to the cube which holds it. Points in each cube are saved into a hush-table. The hush-table is addressed by the cube's indices of three directions x, y, z.

5.2 Building K-neighbour at each point

Dense scattered points are directly employed as the research object. It is essential to research the relationship between each point and its K-neighbour. By K-neighbour location the curvature properties of each point can be reflected.

Based on spatial partition, the index of the cube in which a measured point x_i lies should firstly be calculated. Then, during 27 ($3\times3\times3$) cubes near to x_i K-neighbour are found. Regarding every measured point as a vertex and linking an edge between two points during K-neighbour, a unidirectional map, named Reimann graphy, is developed[11, 12].

5.3 First identification

Due to the zero curvature value of a plane, these plane points can be first extracted from all the measured points in order to improve the efficiency.

1) Calculating tangent plane

We use the same strategy as in Ref.[13] to a calculate a tangent plane. That is, tangent plane P_i at each point x_i is the surface which is the least-square fitting surface of Nbhd (x_i). The tangent plane P_i can be expressed as

$$P_i = S\ (o_i,\ \boldsymbol{n}_i)$$

where o_i is the center of P_i and $\boldsymbol{n}_i$ is the unit normal vector of P_i. o_i can be given as

$$o_i = \frac{1}{k}\sum_{p \ni \mathrm{Nbhd}(x_i)} p$$

For calculating the unit normal vector $\boldsymbol{n}_i$, the matrix

$$\boldsymbol{V}_{\mathrm{C}} = \sum_{p \ni \mathrm{Nbhd}(x_i)} (p - o_i)\ (p - o_i)^{\mathrm{T}}$$

is proposed. The eigenvector corresponding to the minimal eigenvalue of $\boldsymbol{V}_{\mathrm{C}}$ is noted as $\boldsymbol{v}_i$, and $\boldsymbol{n}_i$ is equal to $\boldsymbol{v}_i$. This is proved in Ref.[9]. Since the eigenvector corresponding to the minimal eigenvalue of $\boldsymbol{V}_{\mathrm{C}}$ only needs to be calculated, in this paper the reverse iterative method[13] is proposed to calculate $\boldsymbol{v}_i$.

2) The distance d_i between x_i and the corresponding tangent plane is

$$d_i = |(x_i - o_i)\cdot \boldsymbol{n}_i|$$

3) Marking points at which the underlying surface curvature is very small. A array of visited[i] = false ($i = 0, 2, \cdots, N$) is given. If d_i is smaller than the given threshold, the corresponding visited [i] is equal to ture, where i is the serial number of the 1-D array.

5.4 Second identification

Since curvature tensor of the underlying surface at the remained points of X has a different value at different directions, based on eight fundamental surface types from the signs of K, H and six fundamental types from main curvatures of k_1, k_2[14], we propose K, H at a remained point as the sign of the point. This procedure consists of four steps.

Step 1 If visited[i] = true, the i-th point x_i of X is a point of a plane type and the K, H of it is both zero. Or, go to step 2.

Step 2 We get the K-neighbour of the i-th point of Y. Based on the LBS parameterization method above, the parameter set of $\{(u_i, v_i),\ i = 0, 1, \cdots, k\}$ is obtained. Where (u_0, v_0) is the parameter of the i-th point, and $\{(u_i, v_i),\ i = 1, 2, \cdots, k\}$ are parameters of Nbhd(x_i).

Step 3 Calculate the Gaussian curvature K and the mean curvature H by using the above curvature algorithm.

Step 4 Based on eight fundamental surface types (see Tab. 1) from the signs of K, H and according to K, H of the measured point, the point is labeled into the corresponding surface type. If $i < n$, go to step 1, or this procedure ends.

6 Examples and Discussions

In order to evaluate the usefulness of the developed scheme and the new curvature analysis algorithm, the approach has been applied to a set of sampled data shown in Fig.2. Here $k = 9$.

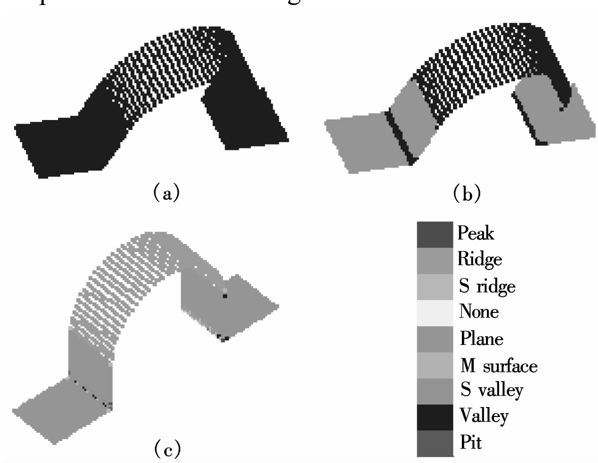


Fig.2 Arch model.(a) Point cloud; (b) Identification of plane points; (c) Segmentation

The point cloud shown in Fig.2(a) consists of 1 887 points sampled from a typical surface of mechanical parts. Based on the sign of K, H at each measured point, Fig.2(b) shows the result of the first identification that extracts points of plane type. The result of the coarse segmentation in Fig.2(c) reflects the fundamental geometric features. Another example shown in Fig.3 consists of 35 875 points sampled by laser range finder ATOS. Coarse segmentation of the model has mostly reflected the corresponding regions.

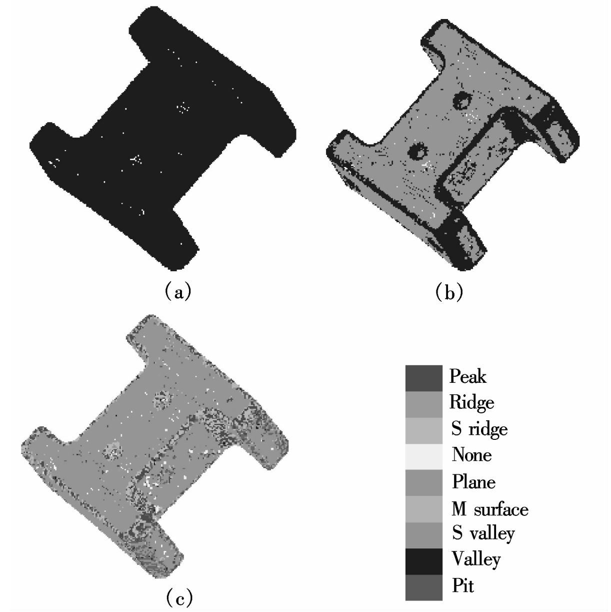


Fig.3 Plane part.(a) Unorganized point cloud; (b) First identification; (c) Segmentation

7 Conclusion

This scheme extends segmentation technology to scattered point cloud without any assumptions. A new LBS parameterization method was applied to the parameters of the i-th point x_i and Nbhd (x_i). Then, the Gaussian and mean curvatures of x_i were estimated by a parametric quadric surface approximation. According to the behavior of surface curvature, the data segmentation was obtained. Spatial partition and K-neighbour techniques used in this method improved the efficiency. Two types of identification were applied also to improve the calculating efficiency. This approach provides the basis of the subsequent reverse engineering steps. Most importantly, the scheme tests the robustness of the new curvature analysis algorithm of randomly measured points. The given examples have shown the essential utility and feasibility of the proposed segmentation method. Future work is going to preprocess the original point data and to refine the coarse segmentation by region growing method for reducing the influence of noise points.

References

[1] Vàrady T, Martin R R, Cox J. Reverse engineering of geometric models — an introduction [J]. *Computer-Aided Design,* **1997, 29**(4): 255 – 268.

[2] Milroy J, Bradley C, Vickers G W. Segmentation of a wrap-around model using an active contour [J]. *Comput-Aided Design,* **1997, 29**(4): 299 – 320.

[3] Yang M, Lee E. Segmentation of measured point data using a parametric quadric surface approximation [J]. *Computer-Aided Design*, **1999, 31**(7): 449 – 457.

[4] Liu S L, Zhou R L, An L L. Region segmentation of a triangle mesh model [J]. *Journal of Nanjing University of Aeronautics and Astronautics* (accepted to appear). (in Chinese)

[5] Huang J, Menq C H. Automatic data segmentation for geometric feature extraction from unorganized 3-D coordinate points [J]. *IEEE Trans on Robo and Auto*, **2001, 17**(3): 268 – 279.

[6] Wang J Y. Segmentation from stratified range image [J]. *Journal of Software,* **2003, 14**(4): 790 – 797.

[7] Liu S X. Reconstruction of curved solids from engineering drawings [J]. *Computer-Aided Design,* **2001, 33** (14): 1059 – 1072.

[8] Ma W Y, Kruth J P. Parameterization of randomly measured points for least squares fitting of B-spline curves and surfaces [J]. *Computer-Aided Design,* 1995, **27**(9): 663 – 675.

[9] Jiang Changjin. *Scientific calculation and C program library* [M]. Hefei: Publishing House of University of Science and Technology of China, 1998.(in Chinese)

[10] Besl P J, Jain R C. Segmentation through variable-order surface fitting [J]. *IEEE Trans Pattern Anal Machine Intell*, **1988, 10**(2): 167 – 192.

[11] Zhou R L, Zhang L Y, Zhou L S. Algorithmic research on surface reconstruction from dense scattered points [J]. *Journal of Software*, **2001, 12** (2): 249 – 255. (in Chinese)

[12] Zhang L Y, Zhou R L, Cai W B, et al. Research on cloud data simplification [J]. *Computer Aided Design & Computer Graph*, **2001, 13** (11): 1019 – 1023. (in Chinese)

[13] Su Xun. Study on the method of surface reconstruction from scattered points in the reverse engineering [D]. Nanjing: College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, 2000.(in Chinese)

[14] Jan Böhm. Towards automated segmentation of dense range scans [J]. *International Archives on Photogrammetry and Remote Sensing IAPRS,* **2002, 34** (5): 137 – 142.

基于一种新的曲率分析算法对散乱数据点云分块

贺美芳　周来水　张丽艳　刘胜兰

(南京航空航天大学 CAD/CAM 工程研究中心, 南京 210016)

摘要: 给出了数据分块系统性方案, 即从仅含有三维坐标的散乱的点云中自动提取几何曲面特性. 首先基于局部基面参数化估算散乱数据点云的局部表面曲率分析是其方案的关键性技术. 再采用由高斯曲率和平均曲率的记号得到的 8 种曲面类型, 就形成初始数据分块. 通过区域增长法可以使粗略数据分块进一步被提取, 得到更小的噪声影响及更精确的区域划分. 其方案得到了实例验证, 具有较强的可操作性和实用性. 基于新曲率算法的分块方案使数据分块技术能够直接运用于散乱数据点云.

关键词: 点云; 分块; 局部基面参数化; 8 种曲面类型; 曲率

中图分类号: TP391.72; TH16