A STUDY OF FEATURE POINTS EXTRACTION BASED ON POINT CLOUD DATA SETS AND MODEL SIMPLIFICATION IN GOAF

Yuanmin Fang*, Jie CHEN, Yonghua XIA

Kunming University of Science and Technology, No.253 Xuefu Road, Kunming, Yunnan Province, P.R.China
TEL: 0871-5140955, FAX: 0871-5153408
E-mail: fangyuanmin@126.com, chenjie3d@126.com, xyh_6603820@126.com

KEY WORDS: Feature Points, Extraction, 3D laser scanner, Point Cloud Data Sets, Model Simplification, Goaf

ABSTRACT:

The shape measurement of goaf is an important job in mine project. With the rapid development of the measuring equipment, 3D laser scanner, which makes the efficiency of traditional work being improved, is widely used in goaf. The final result of 3D laser scan is a huge collection of points. Although different measuring intervals will lead different point cloud density, it also needs a large storage of several MB at least for a single gob area. Point clouds can be used to describe a spatial object's all details with great precision, but there is still a large amount of redundant data. In mathematics, the curvature is the reciprocal of the radius, and a smaller curvature implies a large radius. Intuitively, curvature is the amount by which a geometric object deviates from being flat, or straight in the case of a line; it is defined in different ways depending on the context. Based on the point cloud datasets of a goaf, this thesis offers a new method for modelling the gob area by using the characteristic of curvature of planar curve and the local curvature of spatial curved surface. Actually, the feature points and feature regions of a goaf will be expressed via the curvature parameter. Aiming at simplifying the model, a mathematical model is constructed in order to identify the feature region and non-feature region automatically. The curvature parameter, which is an important controllable variable, can be used to help judging the feature region and non-feature region. Actually, the above-mentioned method could even construct the 3D model of gob area automatically via sampling more points as the feature points in feature region and fewer points in non-feature region. At last, the thesis compares the models that are simplified via different parameters and analyzes how the simplified parameters influence models from the point of system storage space and the extent of distortion etc.

1. INTRODUCTION

As a new technique, 3D laser scanner is widely used in many fields. Actually, the concept of point cloud comes into being accompanying with the appearance of laser scanner. Point cloud consists of a sea of points, in which some otherness exists undoubtedly due to the diversity of various collecting equipments. Detailedly, for all scanners, coordinate is the absolutely necessary information of points; but normal direction, hue and frequency spectrum lie on equipments. Besides, the sampling intervals of different equipments are not the same: some can achieve second-level; the other can only achieve degree-level. In addition, different applications may also lead to different request of point cloud: sampling interval of cultural relic and ancient architecture should be as precise as possible while many redundant statistics still exist in routine mapping works.

When modelling a spatial object, engineers may adopt different measures due to different data-collecting methods. In general way, two measures are mostly used: traditional measure and new-type measure; traditional measure means collecting feature point by total station while new-type measure means gathering point clouds by 3D laser scanner and photographic surveying equipment. Actually, these two measures not only share common ground in object-modelling, which consists of two approaches: data collection and modelling; but also emphasize particularly on different aspects: collecting feature points by total station requires optimized algorithm and emphasizes how to restore genuine model of spatial object with necessary data;

nevertheless, gathering point cloud by 3D laser scanner stresses huge storage space because of its high density scan, which helps obtain the information of spatial object. Nowadays, the research focus of point cloud is feature points collection and model simplification.

When surveying the goaf, scanner, which is manipulated by PDA via wireless communication, extends into the inside through silo. Generally, a scanner drives up or brings down its probe head and scans circumambience with certain intervals. Suppose the interval is 1 degree, the data format is 3D coordinate and each coordinate is stored in the scanner in the form of double float variable (8 bytes long), the necessary storage space is $360*360*3*8/1024/1024=2.9663M$. To sum up, if not compressed, the data size is already nearly 3MB even the interval is at degree-level. If the interval becomes more subtle, the data size will be even larger. Considering a full-scan involving all the goaf of a mine, the data size will be incredibly large.

This thesis takes the rotate-center of scanner motor as the center and establishes orb reference frame (pole reference frame of 3D space) in order to analyze the attributes and features of point clouds; then the thesis offers a way to simplify the feature point extraction and model simplification into math's issue in orb reference frame via data abstraction. Actually, the main contents of this thesis are as follows: (1) definition and solution of feature points; (2) the extraction theory of feature points; (3) the simplifying theory of the model; (4) the precision assessment of the model.

---

* Corresponding author. fangyuanmin@126.com

2. DEFINITION AND SOLUTION OF FEATURE POINTS

Give a planar continuous function $y = f(x)$; suppose c is an inside point of $f(x)$. If the open interval of c exists, and in the interval $f(x) \leq f(c)$ or $f(x) \geq f(c)$ is tenable, then the point $(c, f(c))$ is a local extreme point of function $y = f(x)$. In addition, if $y = f(x)$ is differential at point c, this local extreme point should meet $\frac{dx}{dy} = 0$. In the function graphics, a local extreme point is represented as a critical point of the graphics. In such cases, a local extreme point could be regarded as a feature point (see Figure 1). Furthermore, the infinite situation does not exist in practise, such as a straight line performs as a segment, etc; therefore the endpoint is also included in the scope of critical point.
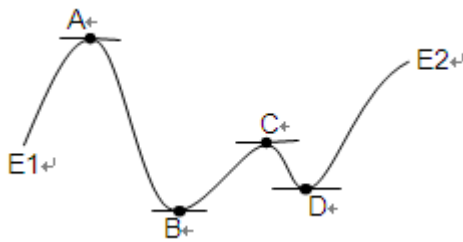

Figure 1: Graphics of function $y = f(x)$
E1, E2 are endpoints; A, B, C, D are critical points

Parameter equation $\left. \begin{array}{l} x = x(t) \\ y = y(t) \end{array} \right\}$ represents planar curve, its

differential coefficient at point t is $\frac{dy}{dx} = \frac{dy}{dt} \Big/ \frac{dx}{dt}$. Actually, there are totally three types of points that meet the demand of critical point: 1. endpoint; 2. the point at which $\frac{dy}{dx} = 0$; 3. the point at which $\frac{dy}{dx}$ does not exist. For instance, in Figure 2, point A and point C meets $\frac{dy}{dx} = 0$; $\frac{dy}{dx}$ does not exist at point B; point E1 and point E2 are endpoints.

As for the implicit function and planar equation, similar conclusion could be drawn, too. The exact definition of feature point is confirmed via summing up the above-mentioned situations:
A point in the domain of smooth planar curve f is a feature point only if:

(1) At this point $f' = 0$;

(2) At this point f' does not exist;

(3) This point is an endpoint.

As can be seen from the definition, feature points could be confirmed via following ways:

(1) At planar smooth curve f, a point at which $f' = 0$ is a feature point;

(2) At planar smooth curve f, a point at which f' does not exist is a feature point.

(3) At planar smooth curve f, endpoints are feature points.
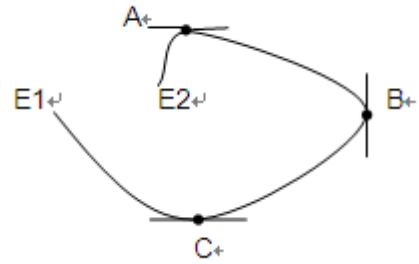

Figure 2: Graphics of parameter equation
E1, E2 are endpoints; A, B, C, D are critical points

3. THE EXTRACTION THEORY OF FEATURE POINTS

After solving out the feature points of a given curve, engineers may rebuild the model via connecting fold line, interpolation or fitting. Actually, the limited number of feature points determines the roughness of the model; in other words, the model is distorted and not sufficiently precise. The only way to build a precise and controllable model is to enlarge the number of feature points.

Viewing intuitively, if there is a planar smooth curve, choosing more points at the segment where the curvature is relatively large and selecting fewer points at the segment where the curvature is small might lead to a better model-construction process. However, the above description contains too many uncertainties, such as how to measure bending, which kind of bending can be described as "relatively large bending or curvature", what is the exact number of "more" in "choosing more points", etc.

In mathematics, curvature is used to measure the bending extent of a curve. If moving along the curve with a certain step, sampling points can be identified. In addition, the exact number of sampling points could be controlled via arc length of the curve.

The parameter equation of planar differentiable function (planar smooth function) $y = f(x)$ is similar to the implicit function. The formula for calculating the curvature is:

$$K(x) = \left| \frac{y''}{\left(1 + y'^2\right)^{\frac{3}{2}}} \right| \qquad (1)$$

The formula for calculating the arc length of curve ab is:

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \, dx \qquad (2)$$

Using parameter equation $\left. \begin{array}{l} x = x(t) \\ y = y(t) \end{array} \right\}, \alpha \leq t \leq \beta$ to represent a curve; suppose $x'(t)$, $y'(t)$ are continuous and not zero in

$[\alpha, \beta]$ at the same time; besides, when increasing from α to β, t passes the curve only once; then the arc length of the curve is:

$$L = \int_{\alpha}^{\beta} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \, dt \qquad (3)$$

Importing curvature threshold K_L and step threshold S_D along the curve as sampling control variables. For $y = f(x)$, taking feature point as sampling center, in the domain of $K \geq K_L$:

$$S = \frac{i \times S_D}{K} \times K_L \qquad (4)$$

Where, K is the curvature at the sampling point; S, which is defined by K, is the step length; i is the sequence number of sampling point.

Via (4), engineers can control the step length at real time. The step length is short at the segment where curvature is large; on the contrary, the step length should be relatively longer at the segment where curvature is smaller. By adjusting S_D, engineers can meet the miscellaneous requests of various curvature radius. Actually, after solving out the step length, the coordinate of sampling point becomes easy to solve with the help of formula (2) and (3).
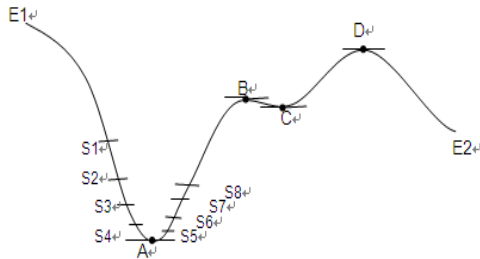

Figure 3: The extraction of sampling point
E1, E2 are endpoints; A, B, C, D are feature points;
S1, S2, S3, S4, S5, S6, S7, S8 are sampling points.

As shown in Figure 3, in the domain of feature point A, eight sampling points (S1 to S8) are collected under the control of curvature threshold K_L and step threshold S_D. Actually, fold line connection, interpolation and fitting can all meet the demand of establishing a good model; besides, smaller-step sampling can be implemented by minimizing K_L or aggrandizing S_D.

4. THE SIMPLIFYING THEORY AND PRECISION ASSESSMENT OF MODEL

The function and parameter equation of models are usually unknown. In order to simplify the model by "the theory of the solution and extraction of feature point", preprocessing the point cloud data is necessary. Taking scanner as the center, establishing pole reference frame and dividing point cloud data

into point sets that lie on separate layers by controlling scan interval; then taking the point at which the vertical axis that crosses through the core of the sphere intersects each layer as center and establishing 2D pole reference frame. For simplifying the model, point sets must be expressed by parametric form. In fact, both interpolation and fitting can be used to establish this "parametric form".

Because of the differences on interpolating basis function, interpolation differs from each other, mainly including polynomial interpolation and spline interpolation etc. The implementation of polynomial interpolation is relatively easy, but Runger phenomenon exists: as the degree of interpolation polynomial increases, severe shake may occur near the interpolation point. However, spline interpolation holds a better effect, but it mainly adopts low-degree polynomial, which might make the construction more complex.

As an approximating measure, fitting usually adopts least square fitting method. The chief problem of fitting is to confirm the form of basis function. Considering that point sets distribute circularly on the layer, engineers may adopt trigonometric function as basis function to construct plane curve.

With the help of interpolation and fitting, point sets on all layers are expressed in parametric form; then the above-mentioned theory could be used to extract feature points. When extraction is finished, models can be realized by triangulation.

Suppose the volume of goaf before simplification is V, the volume after simplification is V_S. The definition of distortion is shown in (5):

$$\varepsilon = \left| \frac{V_S - V}{V} \right| \times 100\% \qquad (5)$$

A general way to calculate the volume is as follow:
(1) Calculating the closed polygon area at all layers, respectively;
(2) Making a "Riemann and" computation; and the result is the volume.

Actually, the volume can also be calculated by integral. It is proper to measure the extent of model distortion by the index "volume loss rate". However, this index can not be used as "absolute measurement". There might also be the case that the value of ε is invalid.

5. CONCLUSIONS

At first, this thesis analyzes the feature of planar curve; then it gives the definition of feature points and offers a method to solve out those points. After that, the thesis imports curvature threshold K_L and step threshold S_D in order to control the precision degree of sampling points. At last, the thesis analyzes the method to simply the data model of point clouds in 3D orb reference frame and 2D pole reference frame; besides, it also imports "model volume loss rate" to measure the extent of model distortion.

REFERENCES

Bajaj, C.L., Bernardini, F., Xu, G., 1995. Automatic reconstruc-

tion of surfaces and scalar fields from 3d scans. In: ACM IGGRAPH 1995, Computer Graphics Proceedings. ACM Press, New York, NY, USA, pp. 109–118.

Bernardini, F., Bajaj, C.L., Chen, J., Schikore, D., 1999. Automatic reconstruction of 3d CAD models from digital scans. *International Journal of Computational Geometry and Applications* 9 (4/5), 327–369.

Edelsbrunner, H., Mucke, E.P., 1994. Three-dimensional alpha shapes. *ACM Transactions on Graphics* 13 (1), 43–72.

Kolluri, R., Shewchuk, J.R., O'Brien, J.F., 2004. Spectral surface reconstruction from noisy point clouds. In: *SGP '04 Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*. ACM Press, New York, NY, USA, pp. 11–21.

Lemmens M. 3D lasermapping. GIM International 2004.18:44–7.

L. Bornaz, F. Rinaudo, Terrestrial Laser Scanning Data Processing, *XXth ISPRS Congress, 12–23 July 2004 Istanbul,* Turkey, Commission, 5, 2004, pp. 514–520.

Lipschutz M. *Theory and problems of di.erential geometry, Schaum Outline Series*. New York: McGraw Hill, ISBN 0 07 0379858.

Mallet, J. L., 2003. Constraining a piecewise linear function defined on a 3d complex (applications to 3d restoration). In: *23rd Gocad Meeting Proceedings*, Nancy, France.

Mallet, J. L., 2002. *Geomodeling*. Oxford University Press, New York, pp. 624.