The Approaching Algorithm of Outline Curves Based on Messy Data *

Dan Liu
Institute of Nautical science & Technology
Dalian Maritime University
Dalian, Liaoning116026, P.R.China
Email:danliu@comp.leeds.ac.uk

Abstract: *The variance of the spatial topological relationships and how to restore these relationships by the geometric relationships between points are researched. A kind of approaching algorithm of outline curves based on messy data is presented. The test results in virtual navigating environment are also given.*

Key words: *Messy data，Approaching，topology*

1.0 Introduction

It plays an important role using curves to interpolate and approach based on messy geometric data in the field of shape designing, data processing and data mining. In every project of engineering, the initial condition to be faced with is discrete data when designing curves. Towards approaching curves, the given information is usually a number of points array. The order feature as well as the approach degree and lubricity should be considered. The senses of order have three aspects of meaning: the first is sufficient lubricity, commonly, with two order lubricity, which means the function has two order continuous derivatives and partial derivatives; the second is that the number of inflexions should be as less as possible; the third is that the variance of curvature should be rather uniform. Towards the spatial curves, the order feature is the order feature of a spatial curve's planar projection on every direction. In the recent years, the automatic synthesis technique of information has come into being, which mainly researches the variances of the spatial points' topological and geometric relationship.
We use the related technique to simulate the target ship and the intruding scenes while the ownship is navigating in the virtual navigating environment.

2.0 The variance of information in the course of approaching shape curves

As we all known，the following factors will affect the geometric relationship of simulating outline curves:

1. The choice of characteristic objects

2. Displacement

3. Distortion

4. Amalgamation

Hence, several special transformations in this course can be obtained:

1. The transformation of single polygon
2. The transformation of two polygons
3. The displacement of polygons
4. The identification and classification of points

3.0 The automatic synthesis technique of information

We must clarify some basic definitions used throughout the work at first. In the topological data architecture, the points are independent; they are connected with each other to compose a line. A line is formed by joining a series of points, begins with the first node, and ends with the last node. A chain is a path on one or more polygons, is also called an arc or a side. Node is the intersection or the end point of lines or chains. A polygon is constituted by an outer ring and null or more inner rings; a ring is composed of one or more chains. Thus it can be seen that the essential geometric architecture of topological data is line, so lines should be synthesized at first. Polishing lines is processed after synthesizing the lines, and then checked the consistency. We will discuss the variance of topological data in the case of several special changes in protraction synthesis.

3.1 The transformation of single polygon

Transform the correlative lines of this polygon. Given a SHAPE polygon $C_i(s) = (u_i(s), v_i(s))$, obtain a polygon S*(s,t) via these lines, t is the parameter of the SHAPE polygon, then we have $S^*(s,t) = (U(s,t), V(s,t))$
Where,

$$U(s,t_i) = u_i(s), V(s,t_i) = v_i(s) \ \forall t \in [0,1]$$

Let S(s,t)=(x(s,t),y(s,t),z(s,t)), thereby

$$S(s,t) = P(t) + a(s,t)\vec{x}(t) + b(s,t)\vec{y}(t)$$

a(s,t) and b(s,t) are defined as follows:
Let

$$m_1 = \min_s U(s,t), m_2 = \max_s U(s,t)$$

$$m_3 = \min_t V(s,t), m_4 = \max_t V(s,t)$$

$$SC_x = \frac{m_2 - m_1}{|l_2(t) - l_1(t)|}, SC_y = \frac{m_4 - m_3}{|l_4(t) - l_3(t)|}$$

Hence

$$a(s,t) = (U(s,t) - m_1)SC_x + < l_1(x) - P_1\ (t), \vec{x}(t) >$$

$$b(s,t) = (V(s,t) - m_3)SC_y + < l_3(x) - P_2\ (t), \vec{y}(t) >$$

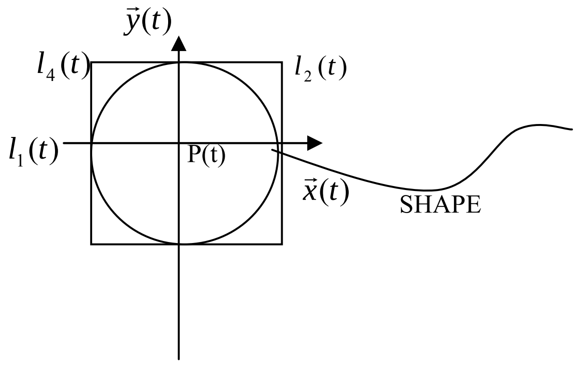


Figure 1 SHAPE curve

3.2 The transformation of two polygons

Transform the inclusive lines of the two polygons.

3.3 The displacement of polygons

process the displacement of involved points in the correlative lines of the polygon, the concrete implement of 3.2 和 3.3 is the same as 3.1.

3.4 The identification and classification of points

Filtrate the dot information through choosing the algorithm. For example, consider the geometric feature of the two classifications. Firstly, compute the expectation $M_1 = \begin{pmatrix} m_{x1} \\ m_{y1} \end{pmatrix}$ of sample ω_1 and the expectation $M_2 = \begin{pmatrix} m_{x2} \\ m_{y2} \end{pmatrix}$ of sample ω_2. Then, calculate the perpendicular bisector between the two points, and use the perpendicular bisector as the criterion function g(X).

The equation of the line is:

$$2x(m_{x1} - m_{x2}) + 2y(m_{y1} - m_{y2}) + m_{x2}^2 - m_{x1}^2 + m_{y2}^2 - m_{y1}^2 = 0$$

When we suppose $P(\omega_1) = P(\omega_2) = 0.5, p(X/\omega_i)$ is the two-dimensional probability density function which obeys $N(\overline{\mu}_i, \sum)$, the previous equation is the exact criterion function of Bayes classification, namely,

$$p(X/\omega_i) = \frac{1}{2\pi |\sum|^{\frac{1}{2}}} \exp\{-\frac{(X-\overline{\mu}_i)^T \sum^{-1}(X-\overline{\mu}_i)}{2}\} (i = 1,2)$$

$$g(X) = W^T X$$

here, $\overline{\mu}_i = M_i = \begin{pmatrix} m_{xi} \\ m_{yi} \end{pmatrix}$ is the expectation of ω_i ,Σis 2x2 covariance matrix,|Σ| is the determinant value of Σ, $X = \begin{pmatrix} x \\ y \end{pmatrix}$ is a training sample, x,y has the uniform variance and is of statistical independence with each other (the correlation function is $\rho = 0$).That is:

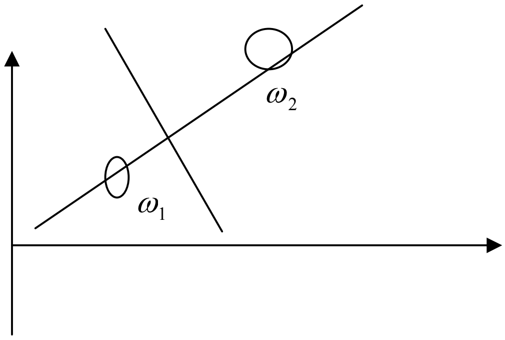


Figure2 The classification of points

$$\sum = \sigma^2 I = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

on account of Bayes classification being the best classification among these situations， so it is reasonable to choose this classification.

Despite the former methods, regarding to the special applications, there are also many other effective methods used in this research work. Fractal theory is an active branch of non-linear Mathematics; it mainly researches the irregular object in the nature and non-linear system. Lots of literatures have forwarded out the applications of Fractal theory in the fields of image texture analysis, natural scene simulation and timing analysis etc.

E.g. in the virtual navigating environment, Fractal theory can be used to process navigating Radar signals, especially the coastline and sea clutters with Fractal features (see Figure 3,4).

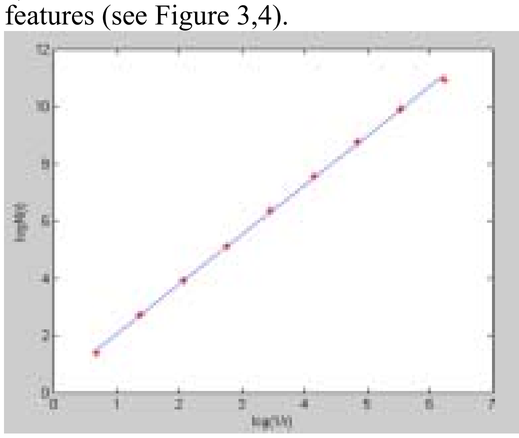
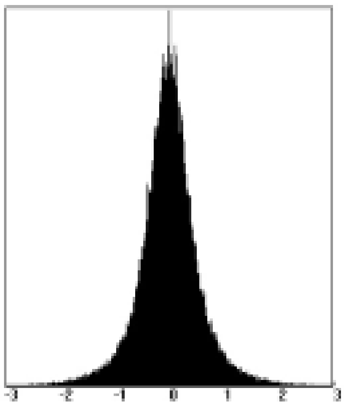
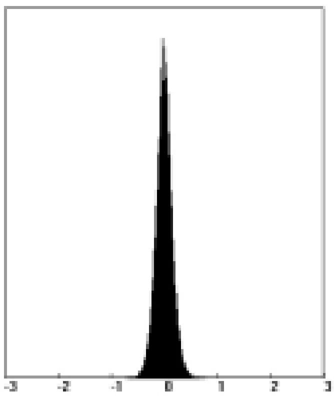


Figure3 The Box dimension (a kind of Fractal dimensions) sketch map of a Navigating Radar signal $f(t)$



(a) High data



(b) low data

Figure4 The histogram of sea clutters increment distribution

Comparing the real sea clutter data with FBM model, the result indicates that both of them match properly. In the research, a Fractal dimension computing program and an algorithm of Fractional Brownian Motion (FBM) curves simulating method is designed.

4.0 Conclusions

Since these methods can preserve the topological relationship of messy data, they can be used to solve the geometric distortion problems effectively.

We can see several better results in virtual navigating environment system using these methods (see Figure 5,6,7,8), includes the application of FBM interpolation method of navigating coastline simulation, target ship simulation and navigating scene simulation .

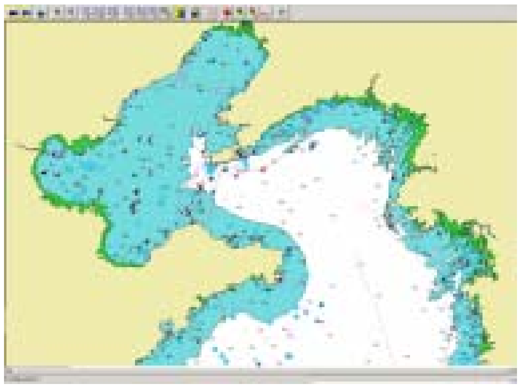


Figure5 Navigating coastline simulation

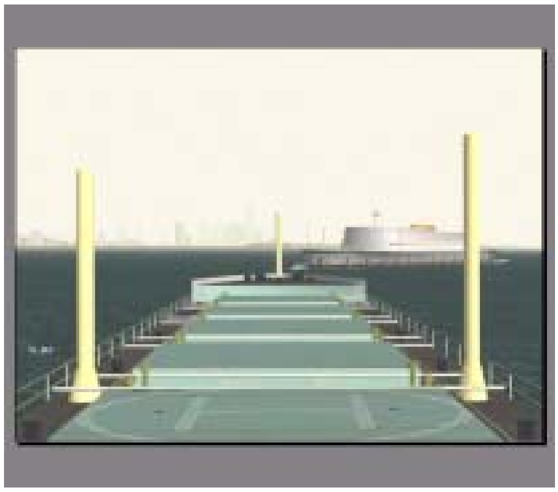


Figure6 Target ship simulation

Figure7 Target ship and navigating scene simulation

Figure8 Navigating scene simulation

References:

[1] A.A.Delucia and R.T.Black.A comprehension approach to automated feature generalization.In Proc.13th Int.Carto.conf.Pages.1978:169-192

[2] H.Edelsbrunner,D.G.Kirkpatrick,and R.Seidel.On the shape of a set of points in the plane. IEEE trans.infor.Theory.IT-29,1983:551-559

[3] Michiel Hagedoorn and Remoc.Velcamp.Metric pattern spaces.Technical Report.UU-CS-1999-03,Utrecht university. Department of computer science, January 1999

[4] Cressie, N., *Statistics for Spatial Data*, Revised Edition, Wiley, New York, 1993

[5] Fayyad, U.M.; Piatestsky-Shapiro G.; Smyth D.; and Uthurusamy R., *Advances in Knowledge Discovery and Data Mining*, Cambridge, MA: AAAI Press/MIT Press, 1996

[6] Fukunaga, K., *Introduction to Statistical Pattern Recognition*, Academic Press, 1990

[7] Groth, R., *Data Mining, A Hands-On Approach for business Professionals*, Prentice Hall, 1998 [with demo software]

[8] Hand, D.J., Mannila, H., Smyth, P., *Principles of Data Mining*, MIT Press, 2000

[9] Jain, A. and Dubes, R., *Algorithms for Clustering Data*, Prentice-Hall, Englewood Cliffs, NJ, 1988

[10] Jordan, M.I., *Learning in Graphical Models*, MIT Press, 1999

[11] Kargupta, Hillol, and Chan, Philip, *Advances in Distributed and Parallel Knowledge Discovery*, MIT/AAAI Press, 2000

---

* This paper is supported by Fok Ying-tong Education Foundation (91077)