Mar. 2006, Vol.21, No.2, pp.261–264 J. Comput. Sci. & Technol.

Determining Knots by Minimizing Energy*

Cai-Ming Zhang[1,2] (张彩明), Hui-Jian Han[1] (韩慧健), and Fuhua Frank Cheng[3] (郑复华)

[1]Department of Computer Science and Technology, Shandong Economic University, Jinan 250014, P.R. China

[2]School of Computer Science and Technology, Shandong University, Jinan 250061, P.R. China

[3]Department of Computer Science, University of Kentucky, Lexington, KY 40506, U.S.A.

E-mail: czhang@sdu.edu.cn

Revised January 18, 2006.

Abstract A new method for determining knots to construct polynomial curves is presented. At each data point, a quadric curve which passes three consecutive points is constructed. The knots for constructing the quadric curve are determined by minimizing the internal strain energy, which can be regarded as a function of the angle. The function of the angle is expanded as a Taylor series with two terms, then the two knot intervals between the three consecutive points are defined by linear expression. Between the two consecutive points, there are two knot intervals, and the combination of the two knot intervals is used to define the final knot interval. A comparison of the new method with several existing methods is included.

Keywords knots, interpolation curve, strain energy, shape preserving

1 Introduction

The problem of constructing parametric interpolating curves is of fundamental importance in computer aided geometric design/modeling, scientific computing and computer graphics. The constructed curve is often required to be smooth and as well as to have the shape suggested by the data points.

The construction of a smooth and visually pleasing parametric interpolating curve requires not only a good interpolation method, but also appropriate choice of the parameter knots. Three methods have been proposed for non-uniform parametrization, namely, chord length method, centripetal model[1] and adjusted chord length method[2,p.111], (referred as Foley's method). Experimental results show that, approximationwise, none of these methods has obvious advantage over the other ones. As far as pleasantness is concerned, centripetal model and Foley's method produce better results than the chord length method. Although these methods are widely used in constructing parametric curves, there are many occasions in which none of these methods can produce a satisfactory result. In those cases, the constructed curves using knots chosen by these methods are obviously different from the shape suggested by the data points. In [3], a new method for determining knots is presented (referred as ZCM method). The knots are determined using a global method. The determined knots can be used to construct interpolants which reproduce parametric quadratic curves if the interpolation scheme reproduces quadratic polynomials.

A new method for choosing knots is presented in this paper. The method first minimizes the internal strain energies of two quadric curves, then the knots are determined by a combination of the two knots corresponding to the two quadric curves. Experiments showed that the curves constructed using the knots by the new method generally has the visually pleasing shape suggested by the given data points.

2 Basic Idea

Let $P_i = (x_i, y_i)$, $1 \leqslant i \leqslant n$, be a set of distinct data points. The goal is to choose a knot t_i for P_i, $i = 1, 2, \ldots, n$. The knot should be chosen so that when used to construct a parametric interpolant $P(t)$ to P_i, $i = 1, 2, \ldots, n$ with the existing methods, $P(t)$ should be with a visually pleasant shape suggested by the data points.

It is supposed that at and near each point P_i, $P(t)$ can be approximated by a curve segment $P_i(t)$ which passes three points P_{i-1}, P_i, P_{i+1}, as shown in Fig.1. Corresponding to $P_i(t)$ and $P_{i+1}(t)$, there are two sets of knots t_i and t_{i+1}, between P_i and P_{i+1}, the knot interval $t_{i+1} - t_i$ is determined by the combination of the two sets of knots.

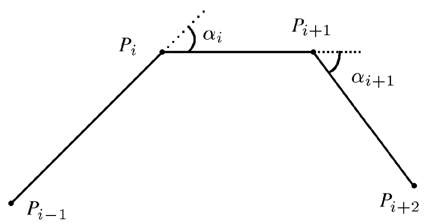

Fig.1. Angle α_i.

For the three knots t_{i-1}, t_i and t_{i+1} corresponding to P_{i-1}, P_i and P_{i+1}, respectively, let

$$\begin{cases} t = t_{i-1} + (t_{i+1} - t_{i-1})s, \\ s_i = (t_i - t_{i-1})/(t_{i+1} - t_{i-1}). \end{cases} \quad (1)$$

Then, the quadric curve $Q_i(t) = (x_i(s), y_i(s))$ that in-

*A preliminary version of this paper appeared in Proc. the 1st Korea-China Joint Conference on Geometric and Visual Computing. Supported by the National Natural Science Foundation of China (Grant Nos. 60403036, 60573114).

terpolates the three points is defined by

$$Q_i(s) = \frac{(s-s_i)(s-1)}{s_i}(P_{i-1}-P_i) + \frac{s(s-s_i)}{1-s_i}(P_{i+1}-P_i) + P_i, \quad (2)$$

where $0 \leqslant s \leqslant 1$, s_i is a variable to be determined between 0 and 1.

There are different methods for choosing knots. The knots can be chosen by the way in which the curve is parameterized or by the way in which the shape of the curve is constructed. It is generally accepted that $Q_i(s)$ would have a desirable shape if it has the minimum internal strain energy. The idea here is to choose an s_i that would make $Q_i(s)$ have the minimum internal strain energy, and hence make $Q_i(s)$ have a desirable shape. Thus s_i is determined by minimizing the following objective function:

$$G(s_i) = \int \Big[\frac{\mathrm{d}^2\kappa(\varpi)}{\mathrm{d}\varpi^2}\Big]^2 \mathrm{d}\varpi, \quad (3)$$

where $\kappa(\varpi)$ is the curvature of $Q_i(s)$, defined as a function of the arc length of $Q_i(s)$.

If the magnitude of the first derivative of $Q_i(s)$ is close to a constant or the arc length of $Q_i(s)$ is chosen as its parameter, then the objective function (3) can be approximated by

$$G(s_i) = \frac{1}{C}\int_0^1 |Q_i''(s)|^2 \mathrm{d}s.$$

where C is a constant.

Straightforward computation shows that

$$G(s_i) = \frac{4}{C}\Big(\frac{d_{i-1}^2}{s_i^2} - \frac{2\cos\alpha_i}{s_i(1-s_i)}d_{i-1}d_i + \frac{d_i^2}{(1-s_i)^2}\Big),$$

where d_i denotes the distance from P_i to P_{i+1}, and α_i the angle between the vectors $\boldsymbol{P}_{i-1}\boldsymbol{P}_i$ and $\boldsymbol{P}_i\boldsymbol{P}_{i+1}$, as shown in Fig.1.

The free variable s_i is determined by

$$\frac{\mathrm{d}G(s_i)}{\mathrm{d}s_i} = 0. \quad (4)$$

The solution of (4) is a function of α_i. For special cases, we have the following Theorem.

Theorem. *For $\alpha_i = 0$, $\alpha_i = \pi/2$ and $\alpha_i = \pi$, setting*

$$s_i = \frac{d_{i-1}}{d_{i-1}+d_i}, \quad (5)$$

$$s_i = \frac{d_{i-1}^{2/3}}{d_{i-1}^{2/3}+d_i^{2/3}}, \quad (6)$$

$$s_i = \frac{d_{i-1}^{1/2}}{d_{i-1}^{1/2}+d_i^{1/2}}, \quad (7)$$

respectively, makes the objective function (4) *reach its minimum values.*

If P_{i-1}, P_i and P_{i+1} are collinear, and s_i is defined by (5), then Q_i is a straight line.

Substituting (7) into $Q_i(s)$ (2), one gets $|\mathrm{d}Q_i(s_i)/\mathrm{d}s| = 0$. This means that if vectors $\boldsymbol{P}_{i-1}\boldsymbol{P}_i$ and $\boldsymbol{P}_i\boldsymbol{P}_{i+1}$ are of opposite direction, and s_i is defined by (7), then $|\mathrm{d}Q_i(s_i)/\mathrm{d}s| = 0$. If $Q_i(s)$ is viewed as the trajectory of a particle, $|\mathrm{d}Q_i(s_i)/\mathrm{d}s| = 0$ means that the speed of the particle is zero when passing through the point P_i. The zero speed makes it easy for the particle to turn around at P_i.

3 Determining Knot Interval $t_{i+1} - t_i$

Based on (5)–(7), s_i can be expressed as a function of α_i as follows

$$s_i = \frac{d_{i-1}^{f(\alpha_i)}}{d_{i-1}^{f(\alpha_i)} + d_i^{f(\alpha_i)}}$$

where $f(0) = 1$, $f(\pi/2) = 2/3$ and $f(\pi) = 1/2$.

It follows from (1) that $t_{i+1} - t_i$ can also be regarded as a function of α_i, i.e., $t_{i+1} - t_i = t(\alpha_i)$, then $t_{i+1} - t_i$ can be expanded as a Taylor series at $\alpha_i = 0$

$$t_{i+1} - t_i = t(0) + \frac{\mathrm{d}t(0)}{\mathrm{d}\alpha_i}\alpha_i + O(\alpha_i^2). \quad (8)$$

For $\alpha_i = 0$, setting $t_i - t_{i-1} = d_{i-1}$ and $t_{i+1} - t_i = d_i$ will make $P_i(s)$ being a straight line, this is a natural choice. Thus we have $t(0) = d_i$. Since $t_{i+1} - t_i = d_i$ and $t_i - t_{i-1} = d_{i-1}$ at $\alpha_i = 0$, it follows from (8) that, for $0 \leqslant \alpha_i \leqslant \pi$, $t_i - t_{i-1}$ and $t_{i+1} - t_i$ can be approximated by the first and second terms of their Taylor expansions as follows:

$$\begin{cases} t_i - t_{i-1} = d_{i-1}(1 + b_i\alpha_i/\pi), \\ t_{i+1} - t_i = d_i(1 + a_i\alpha_i/\pi). \end{cases} \quad (9)$$

where b_i and a_i are to be determined. Note that (9) satisfies (5) for $\alpha_i = 0$.

Based on (9), (1) can be written as

$$\frac{d_{i-1}(\pi + b_i\alpha_i)}{d_{i-1}(\pi + b_i\alpha_i) + d_i(\pi + a_i\alpha_i)} = s_i \quad (10)$$

It follows from (10) and Theorem that the values of a_i and b_i can be approximated by the following equations

$$\frac{d_{i-1}(1+b_i/2)}{d_{i-1}(1+b_i/2) + d_i(1+a_i/2)} = \frac{d_{i-1}^{\frac{2}{3}}}{d_{i-1}^{\frac{2}{3}} + d_i^{\frac{2}{3}}},$$

$$\frac{d_{i-1}(1+b_i)}{d_{i-1}(1+b_i) + d_i(1+a_i)} = \frac{d_{i-1}^{\frac{1}{2}}}{d_{i-1}^{\frac{1}{2}} + d_i^{\frac{1}{2}}}.$$

The solutions are

$$\begin{cases} b_i = \left(\frac{d_i}{d_{i-1}}\right)^{\frac{1}{3}} + \left(\frac{d_i}{d_{i-1}}\right)^{\frac{1}{6}} - 1, \\ a_i = \left(\frac{d_{i-1}}{d_i}\right)^{\frac{1}{3}} + \left(\frac{d_{i-1}}{d_i}\right)^{\frac{1}{6}} - 1. \end{cases} \tag{11}$$

We now discuss how to determine the value of $t_{i+1} - t_i$. From the second case in (9), one gets an approximated value of $t_{i+1} - t_i$. By working with α_{i+1} using P_i, P_{i+1} and P_{i+2}, one gets another approximated value of $t_{i+1} - t_i$. Considering the second term of the Taylor series, $t_{i+1} - t_i$ can be written as the following more general forms

$$\begin{cases} t_{i+1} - t_i = d_i(1 + \mu_i), \\ t_{i+1} - t_i = d_i(1 + \lambda_i), \end{cases} \tag{12}$$

where $\mu_i = (1 + A_i)a_i\alpha_i$, $\lambda_i = (1 + B_i)b_{i+1}\alpha_{i+1}$, with A_i and B_i being unknown.

These two expressions can be used to define $t_{i+1} - t_i$. Since the formation of $t_{i+1} - t_i$ is related to both α_i and α_{i+1}, $t_{i+1} - t_i$ should be defined as a combination of these two expressions.

$$t_{i+1} - t_i = d_i(1 + w_i\mu_i + (1 - w_i)\lambda_i), \tag{13}$$

where $t_1 = \alpha_1 = \alpha_n = d_0 = d_{n+1} = 0$,

$$w_i = \frac{d_{i-1} + d_i}{d_{i-1} + 2d_i + d_{i+1}} \tag{14}$$

is a weight function. (13) holds for $i = 1, 2, \ldots, n-1$.

By testing many sets of data points, the following values of A_i and B_i has been proven to be a better choice to define $t_{i+1} - t_i$

$$\begin{cases} A_i = \rho d_{i-1}/d_i, \\ B_i = \rho d_{i+1}/d_i \end{cases} \tag{15}$$

with $1 \leqslant \rho \leqslant 2$ being a parameter to control the shape of the curve (in general case, $\rho = 1$).

4 Discussion

For evenly spaced data points, knots defined by (13) generally produce better results when applied to cubic spline interpolation. Fig.2 are several examples. The data points used to produce the curves are defined by

$$\begin{cases} P_3 = (0, 0), \quad P_4 = (1, 0), \\ P_1 = P_3 + (-2\cos(\alpha_3), 2\sin(\alpha_3)), \\ P_2 = (P_1 + \omega P_3)/(1 + \omega), \\ P_6 = P_4 + (2\cos(\alpha_4), 2\sin(\alpha_4)), \\ P_5 = (P_4/\omega + P_6)/(1/\omega + 1). \end{cases} \tag{16}$$

where $\omega > 0$.

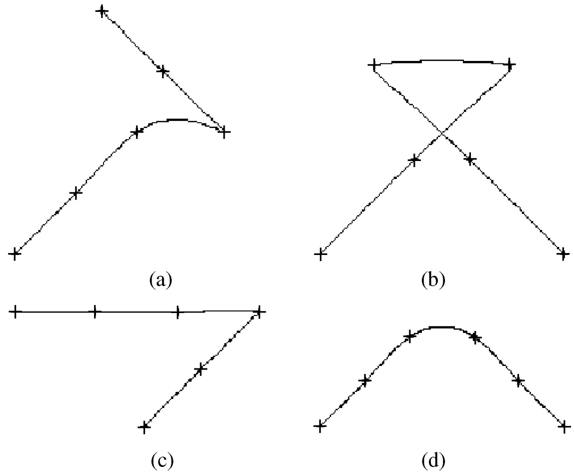

Fig.2. (a) $\alpha_3 = -3\pi/4$, $\alpha_4 = 3\pi/4$. (b) $\alpha_3 = -\pi/4$, $\alpha_4 = -3\pi/4$. (c) $\alpha_3 = \Pi$, $\alpha_4 = -3\pi/4$. (d) $\alpha_3 = -3\pi/4$, $\alpha_4 = -\pi/4$.

The curves shown in Fig.2 are cubic spline curves interpolating the data points (marked by symbol "+") defined by (16) by taking $\omega = 1$ and varying α_3 and α_4.

The experiments show that the knots by (13) produce satisfactory cubic spline curves for ω, α_3 and α_4 satisfying $0 \leqslant \omega \leqslant 1.5$, $0 \leqslant \alpha_3 \leqslant 3\pi/4$ and $-3\pi/4 \leqslant \alpha_4 \leqslant 3\pi/4$.

For unevenly spaced data points, the values of some d_{i-1}/d_i and d_{i+1}/d_i could become larger and, consequently, lead to large $t_{i+1} - t_i$ and unsatisfactory results. This shortcoming can be overcome by redefining $t_{i+1} - t_i$ as follows.

$$t_{i+1} - t_i = d_i(1 + w_i\bar{\mu}_i + (1 - w_i)\bar{\lambda}_i), \tag{17}$$

$$\bar{\mu}_i = \begin{cases} \mu_i, & \text{if } \mu_i \leqslant \kappa_i, \\ \kappa_i + \dfrac{\mu_i - \kappa_i}{1 + \mu_i - \kappa_i}, & \text{otherwise} \end{cases}$$

$$\bar{\lambda}_i = \begin{cases} \lambda_i, & \text{if } \lambda_i \leqslant \kappa_i, \\ \kappa_i + \dfrac{\lambda_i - \kappa_i}{1 + \lambda_i - \kappa_i}, & \text{otherwise} \end{cases}$$

where μ_i and λ_i are defined by (12), and

$$\kappa_i = \frac{d_{i-1}}{d_{i-1} + d_i} + \frac{d_{i+1}}{d_i + d_{i+1}} + \sqrt{2\alpha_i\alpha_{i+1}}.$$

5 Experiments

Three sets of representative data points have been used to compare the new method with the centripetal, Foley's and ZCM methods. The comparison is performed using knots determined by these methods in the construction of a parametric cubic spline curve which interpolates the given data points.

The three-point difference formula is used to determine the end conditions of the spline curve. The three sets of representative data points are Akima[5], FRN 15A[4] and Brodlie[6] data points. The curves generated by these methods are shown in Figs. 3 and 4, where the symbol "+" denotes the location of a given data point, figures (a), (b), (c) and (d) are produced by the

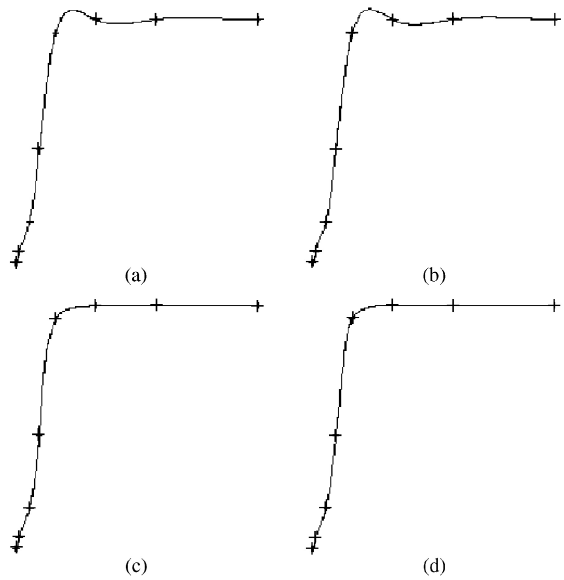

Fig.3. FRN 15A[4] data points.

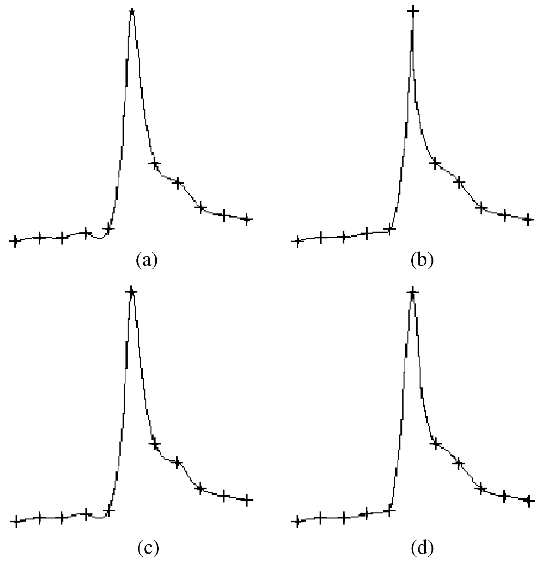

Fig.4. Brodlie[6] data points.

centripetal model, Foley's, ZCM and new methods, respectively. The result for Akima[5] data are not shown as it is similar to the one shown in Fig.3.

6 Conclusions

A new method of choosing knots for parametric polynomial interpolation process is presented. The knots are determined by minimizing the internal strain energy of an interpolation curve. Experiments also show that the curves constructed using the knots by the new method generally has the visually pleasing shape suggested by the given data points.

It should be pointed out that for the curve that changes rapidly, its strain energy cannot be approximated by a simple objective function. In this case, the method presented in this paper might not give better results. This problem will be our future research work.

References

[1] Lee E T Y. Choosing nodes in parametric curve interpolation. *CAD,* 1989, 21(6): 363–370.
[2] Farin G. Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide. Academic Press, 1988.
[3] Zhang C, Cheng F, Miura K. A method for determing knots in parametric curve interpolation. *CAGD,* 1998, 15: 399–416.
[4] Fritsch F N, Carlson R E. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.,* 1980, 17: 238–246.
[5] Akima H. A new method of interpolation and smooth curve fitting based on local procedures. *J. ACM,* 1970, 17(4): 589–602.
[6] Brodlie K W (ed.). A Review of Methods for Curve and Function Drawing. Mathematical Methods in Computer Graphics and Design, London, UK: Academic Press, 1980, pp.1–37.

Cai-Ming Zhang is a professor and Ph.D. supervisor in the School of Computer Science and Technology at Shandong University. He is now also the dean and professor of the School of Computer Science and Technology at the Shandong Economic University. He received the B.S. and M.E. degrees in computer science from Shandong University in 1982 and 1984, respectively, and the Dr. Eng. degree in computer science from the Tokyo Institute of Technology, Japan, in 1994. From 1997 to 2000, Dr. Zhang has held visiting position at the University of Kentucky, USA. His research interests include CAGD, CG, information visualization and medical image processing.

Hui-Jian Han is an associate professor in the Computer School of Shandong Economic University. He received the M.E. degree in computer science from Shandong University in 2003. His current research interests include computer graphics, information visualization and CAGD.

Fuhua Frank Cheng is a professor of computer science and supervisor of the Graphics and Geometric Modeling Lab at the University of Kentucky where he joined the faculty in 1986. He received the B.S. and M.S. degrees in mathematics from the "National Tsing Hua University" in 1973 and 1975, respectively, the M.S. degree in mathematics, the M.S. degree in computer science, and the Ph.D. degeree in applied mathematics and computer science from the Ohio State Unversity, in 1978, 1980 and 1982, repectively. Dr. Cheng has held visiting positions at the University of Tokyo and the Unversity of Aizu, Japan. His research interests include computer aided geometric modeling, computer graphics, and parallel computing in geometric modeling and computer graphics.