SELECTION OF THE NUMBER OF CONTROL POINTS FOR SPLINE SURFACE APPROXIMATION

Nirant V. Puntambekar
Andrei G. Jablokow

Department of Mechanical Engineering
Pennsylvania State University
University Park, PA, USA

ABSTRACT

Parametric spline surfaces represent an important surface type in the process of reverse engineering. Various surface fitting techniques are available for fitting of splines to sets of data points, here we consider a least squares fitting technique. A user input to this algorithm is the number of control points to be used for the fitting in the two parametric directions. Using a larger number of points, gives a more accurate fitted surface but also results in a surface with a number of undulations or surface wiggles which may not be desired. This paper is concerned with the optimal selection of the number of control points to be used for the least squares surface fitting of B spline surface patches. An optimal method based on multicriteria optimization is presented to decide on the number of control points to be used in the fitting.

Keywords: Geometric modeling, Reverse engineering, Surface fitting, Splines.

INTRODUCTION

Reverse engineering is an important and emerging process which deals with the generation of solid models from point data from objects. This methodology is used in applications such as manufacturing parts designed from model testing or by artistic designers, replicating antiques, redesign of existing components without existing drawings, etc. typically in fields like automobile, ship building, aerospace and manufacturing industry. Point data on the surfaces of the objects is first obtained using measuring equipment such as CMM's, laser scanners, etc. which is then processed to generate a model of the object. A widely used approach for model generation is to identify faces/regions on the surface of the object and then fit surfaces to the points representing the same. These fitted surfaces can then be intersected or patched together in the proper topological order to generate the solid model [Punta94] which can then be used for a variety of applications in a computer aided design and manufacturing system.

An important step in this reverse engineering procedure is the fitting of surfaces to the data points. The surfaces can be algebraic or parametric depending on the underlying surface. Parametric surface fitting using B-spline or NURBS surfaces represents a versatile method of representing the face/region of a free form surface. Least squares algorithms are available for the fitting of B-spline/NURBS surfaces to a set of data points of rectangular topology [Piegl91]. A user input for these algorithms is the degree of the basis spline functions and the number of control points to be used for the surface. Splines of degree three are most commonly used for efficient computation [Roger90] and so the number of control points to be used is an important parameter affecting the final fitted surface.

Using a large number of control points (equal to the number of data points in the limit) makes the surface approach an interpolating surface (passing through all the data points, hence would have a zero error). Such a surface would be wiggly (not fair) as it is forced to interpolate the data points. On the other hand, using a fewer number of control points would produce a flatter(fairer) surface but would have a larger error associated with it. This is illustrated in Figure A.

Interpolating Surface
- *low error*
- *more undulations*
- *more control points*

Approximating Surface
- *high error*
- *fewer undulations*
- *less control points*

Figure A: Interpolating Vs. Approximating Surface

Hence, proper selection of the same, results in a surface with a low error and having a smooth and aesthetic appearance (fair surface). This paper is concerned with the issue of the selection of the number of control points to be used in the fitting of B-spline/NURBS surfaces to a set of data points. A scheme for the optimal selection of the same is presented.

SURFACE APPROXIMATION

Given a set of data points, there are two ways one can obtain a surface which represents the given set of points - interpolation and approximation. In interpolation, a surface which passes through each of the data points is obtained whereas in approximation, a surface which need not pass through any of the data points but best represents the given data set in an average sense is desired. A variety of error norms can be minimized in the approximation, the most common being the least squares approximation in which the sum of the squares of the errors (between a data point and the corresponding point on the surface) is minimized. We shall be using a least squares algorithm for the fitting of the surfaces.

Piegl [Piegl91] describes an algorithm for the fitting of a tensor product B-spline surface patch to a gridded set of data using a least squares technique for the approximation. Consider a given set of $(n_1+1) \times (m_1+1)$ weighted data points $Q_{r,s}$, where $r = 0,\ldots, n_1$ and $s = 0,\ldots, m_1$. It is desired to find a degree (p, q) surface that agrees as far as possible with $Q_{r,s}$. That is,

$$Q_{r,s} = S(u_r, v_s) = \sum_{i=0}^{n}\sum_{j=0}^{m} P_{i,j}N_{i,p}(u_r)N_{j,q}(v_s) \quad (1)$$

where $P_{i,j}$ are the unknown control points for the patch and $N(u)$ and $N(v)$ are the standard B-spline blending functions in the two parametric directions, u and v. The above equation can be evaluated for each of the data points and this gives us a set of equations which can be assembled in matrix form as

$$Q = NP \quad (2)$$

where,

$$N = N_{i,p}(u_r)N_{j,q}(v_s) \quad (3)$$

Here Q is a $(n_1+1)(m_1+1)\times 3$ matrix of data points, P is the $(n+1)(m+1)\times 3$ matrix of control points and N is a $(n_1+1)(m_1+1)\times(n+1)(m+1)$ matrix of blending functions. Equation 2 represents an overdetermined system in surface approximation as there are more data points than control points. To solve this equation, one needs the parameter values at which the data points are assumed and the appropriate knot vectors. The weights are assumed to be unity for simplicity. The parameter values in each of the parametric directions are computed using the centripetal method [Lee89] and the knot vectors are obtained from these using an averaging scheme [Piegl91]

The degree of the surface and the number of control points to be used for the patch in the two parametric directions is an user input for this fitting algorithm and therefore needs to be specified beforehand. Equation 2 can be solved for the unknown control points P by taking the generalized pseudo-inverse as

$$P = (N^T N)^{-1} N^T Q \quad (4)$$

The control points along with the knot vectors define the B-spline patch which has been fitted to the given set of data points in a least squares sense.

SURFACE CHARACTERISTICS

The fitted parametric surface obtained from the above least squares approximation algorithm will be a surface which represents the given data points in an average sense. There are two important characteristics which will determine the quality of this surface - the error of fitting in the surface and the fairness of the surface. Both of these are dependant on the number of control points used for the fitting. These factors are now discussed below.

Error: Since we are using data approximation, the fitted surface may not pass through any of the data points and there is always an error associated with the surface. This error for the surface can be measured as the sum of the errors between each of the data points and the corresponding point on the surface which is nearest to the data point (the surface evaluated at the parameter values assigned to that data point during the approximation). Thus the error norm can be specified as

$$e = \sum_{datapoints} dist(P_{data} - P_{surf}) \quad (5)$$

where P_{data} is a data points and P_{surf} is a point on the surface closest to the data point.

This error in the surface is dependant on the number of control points used for the fitting and it is observed that it reduces as we utilize a higher number of control points for the surface. In the limit when we choose the number of control points equal to the number of data points, the surface will interpolate all the data points resulting in a zero error of fitting.

Fairness: Fairness or smoothness of the surface is a characteristic which measures the aesthetic nature of the surface. Thus, the fairer the surface, the lesser the number of undulations or wiggles it has, making it more pleasing in appearance. Fairness is often a criteria used by surface designers to evaluate whether the objects they are designing are artful as well as functional. To give a physical analogy, a curve can be considered fair if it can be drawn using a small number of french curves [Farin87]. Curvature is a quantity which is intrinsically linked to the fairness of the curve/surface. Mathematically, fairness implies that the curve/surface has a smooth curvature variation without any discontinuities or sign changes.

A number of schemes have been used by researchers to measure the fairness of surfaces. Munchmeyer [Munch78] proposes a network of lines of curvature as a tool to analyze the fairness of a surface which is subsequently used in the smooth outer surface design of high performance ships and aircraft. Izumida [Izumi79] uses an interactive technique based on equi-gaussian curvature lines on the surface to device a system for ship hull form definition and design. Farin [Farin89] makes use of curvature plots in a curve fairing algorithm for improving the fairness of designed curves/surfaces. Points on the curvature plot having the greatest slope discontinuity are recognized and the curve is locally adjusted and faired in that region. Another technique of interrogating the fairness is the method using reflection lines [Klass80] which are the patterns formed on the surface by the mirror images of a number of parallel fluorescent light strips.

Most of the above techniques are interactive in the sense that the user decides on the fairness after visually evaluating the plots on the screen. For our purposes, we will be using a non-interactive global measure of fairness based on the gaussian curvature evaluated at different points which will give us an estimate of the fairness associated with the surface.

At any point on a surface, there exist two mutually orthogonal directions along the surface such that the curvature values along these directions are extreme values. The curvature evaluated along any other direction at that point will always lie in-between these two values. These two curvature values are known as the principal curvatures, κ_{min} and κ_{max} and the two directions are referred to as the principal directions. The product of the two curvatures is known as the Gaussian curvature at that point.

$$K_g = \kappa_{min}\kappa_{max} \tag{6}$$

This gaussian curvature serves to characterize the local shape of the surface [Roger90], a negative value implies a hyperbolic region whereas a positive value implies an elliptic region. If the Gaussian curvature value is zero at all points on a surface, it implies that the surface is developable can be unfolded into a plane without any stretching or tearing. Thus, changes in this local behaviour from elliptical to hyperbolic correspond to undulations in the surface, serving as a measure of fairness.

We introduce a new coefficient called the wiggle coefficient, which will measure the global fairness value for a surface. The gaussian curvature is evaluated at a number of equispaced points (parametrically) on the surface, more specifically at equispaced points along the curves obtained by keeping one of the parameters constant and varying the other. This is repeated for all the curves along the surface. The wiggle coefficient, W is then defined as the number of times the value of gaussian curvature changes in sign over the surface.

$$W = \sum_{surface} signchanges(K_g) \tag{7}$$

This coefficient will be used as an estimate of the fairness of the surface. Thus lower the value of the wiggle coefficient, fairer or smoother will be the fitted surface.With regards to parametric surface fitting, the fairness of the resulting surface is a function of the number of control points used. Higher the number of control points less fair would the surface be, with a number of undulations on the surface.

MULTICRITERIA OPTIMIZATION.

This is a technique used for the constrained optimization of a number of objective functions. The aim of this multicriteria optimization is to try to find the best compromise solution to all the given objective functions. A number of techniques are available for the same and we will be using the Global Criteria Method [Jendo85]. The general multicriteria optimization problem can be stated as follows.

Minimize the objective functions

$$f_j(x) \quad , \quad j = 1, 2, \ldots, k \tag{8}$$

subject to the constraints

$$g_i(x) \le 0 \quad , \quad i = 1, 2, \ldots, m \tag{9}$$

Here x is the vector of design variables, given by

$$x^T = \{x_1, x_2, \ldots, x_n\} \tag{10}$$

The solution to this problem is obtained by formulating a global criteria based on the objective functions subject to the constraints as,

$$F^{(p)} = \left(\sum_{j=1}^{k} |\tilde{f}_j(x)|^p\right)^{1/p} \tag{11}$$

where $\tilde{f}_j(x)$ are the normalized objective functions which have the form

$$\tilde{f}_i(x) = \frac{f_i(x) - min(f_i(x))}{max(f_i(x)) - min(f_i(x))} \qquad (12)$$

and p is a parameter whose value can be from $1 \le p \le \infty$, usually taken as 2 [Jendo85]

Thus the multicriteria optimization problem is transformed into a standard constrained optimization problem involving a single objective function (Equation 11). This problem can be solved by a number of standard techniques available such as the simplex method, box algorithm, golden search method (one dimensional problems), etc. [Haftk92].

APPROACH

As noted in the above section the two parameters characterizing the fitted surface are the error and the fairness. In surface fitting, both of these are functions of the number of control points used. The two parameters, the error using the error norm defined by Equation 5 and the fairness using the wiggle coefficient given by Equation 7 will be evaluated for a given surface using different values of the number of control points during fitting. Second degree functions are then obtained representing these two sets of data values using a least squares technique. These two functions will then be simultaneously minimized with the number of control points in the two directions as the variables, using the multi criteria optimization method. Thus the optimum values for the number of control points in the two directions can be obtained which minimizes both the error and wiggle coefficient (maximizes the fairness) for the surface. This procedure can be summarized as

1. Obtain error(Equation 5) and fairness (Equation 7) values after least squares fitting of surface (using Equation 4) with different number of control points.
2. Obtain error and fairness functions (least squares) (f1 and f2 in Equation 8) which best represents each of these criteria as a function of the number of control points.
3. Optimize these individual criteria using multicriteria optimization (Equation 11) to obtain the value of the optimum number of control points to be used in the surface fitting.

We now illustrate the above methodology with the help of a few examples. The first example is that of a rectangular planar patch obtained by fitting a B-spline surface to a set of gridded data points with random perturbations (maximum amplitude 0.1) added to the coordinates to simulate the errors introduced by measuring equipments. Equal number of control points are used in the two parametric directions for the fitting. The error and the wiggle coefficient are plotted against the number of control points as shown in Figure 1. It is seen that as the number of control points used increase, the error in the surface reduces whereas the value of the wiggle coefficient increases (implying a reduction in the fairness for the surface). Polynomial functions are fitted to each of these data sets and the global function F is then formulated as given by Equation 11 which is then minimized with respect to the number of control points. The plot of this cost function is as shown in Figure 1.

Similar analysis is performed for a curved patch (Figure 2) and a circular sector patch with uneven spacing of the data points (Figure 3) and the results are as shown below.

It is seen that the optimum value of the number of control points to be used for the fitting, which minimizes the error and maximizes the fairness for all of the above three cases is close to 6.0 (The nearest integer value). In order to further verify this, a similar analysis was performed for the surfaces shown in Figure 4 and the optimum values are computed for the number of control points, tabulated as shown in Table 1.

Surface	**Optimum no. of Cont. Pts.**
Hat	6.12
Mobius Strip	6.16
Monkey Saddle	6.65
Doubly Curved Surface	6.28

TABLE 1. Optimum Values for other surfaces

To test the effect of the input data point density on the selection of the number of control points, various point grids of different densities were generated for the doubly curved surface. The optimizing procedure was carried out for each of the different sets of data points generated. The optimum values for the number of control points are as listed in Table 2.

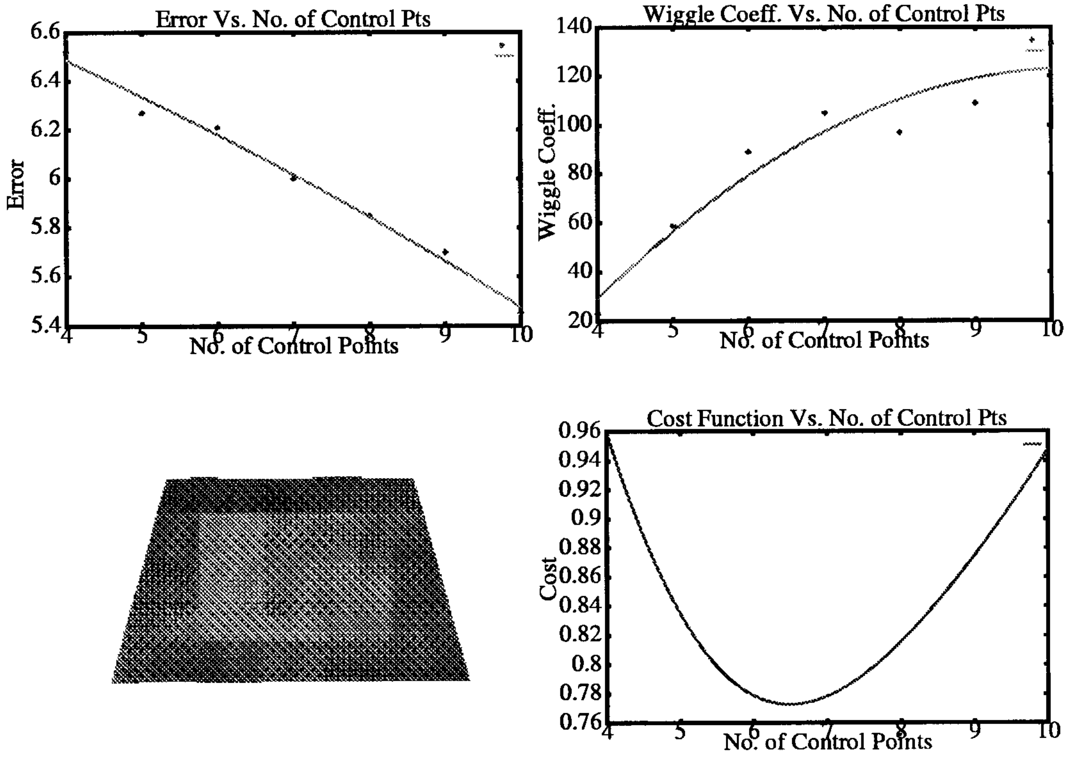


FIGURE 1. Error, Fairness and Cost function for plane

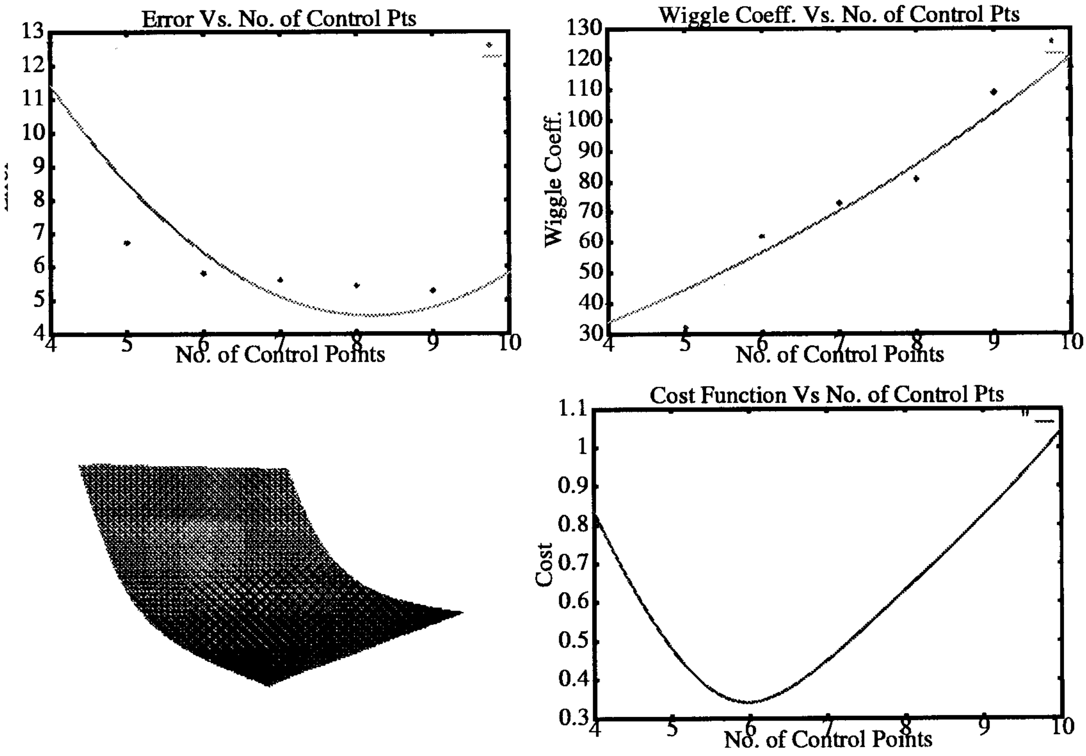


FIGURE 2. Error, Fairness and Cost Function for a curved patch

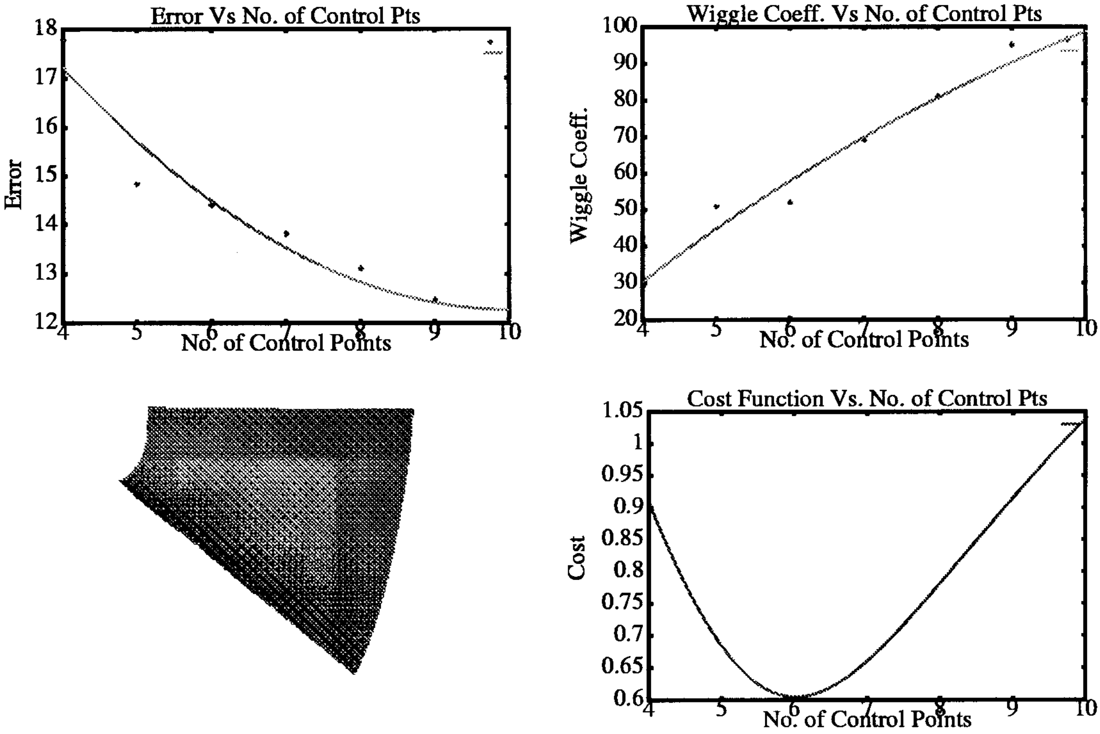
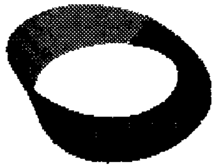
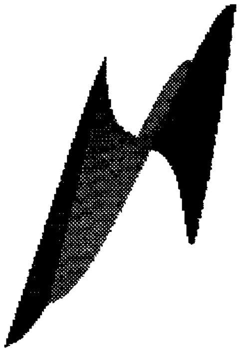


FIGURE 3. Error, Fairness and Cost Function for a circular sector patch

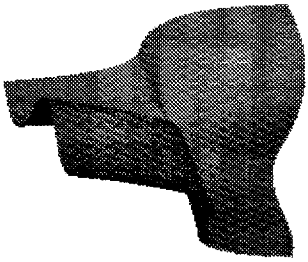
a) Hat



b) Mobius Strip



c) Monkey Saddle



d) Doubly Curved Surface

FIGURE 4. Other Surfaces

Grid Density	Optimum No. of Control Pts
15 x 15	6.184
21 x 21	6.282
31 x 31	5.913
41 x 41	6.399
45 x 45	6.474

TABLE 2. Effect of grid density on the control point selection

Unequal Number of Control Points: The above procedure can be extended for the case of unequal number of control points in the two parametric directions. In this case we have two variables for the optimization, m and n, which are the number of control points in the two parametric directions. The error in the surface and the fairness values can be calculated for different values of m and n and quadric polynomial surfaces can be fitted to these data sets. These fitted surfaces will be quadratic in m and n and these can be used to formulate the global objective function given by Equation 11. This objective function can be minimized for m and n using standard multivariable optimization. The values obtained represent the optimal combination of the control points which when used in the surface fitting will result in a surface with optimal error and fairness characteristics.

DISCUSSION AND CONCLUSIONS

Parametric spline surfaces are an important surface type in the reverse engineering process for the reconstruction of free form and nonuniform surfaces.This paper has presented a method to optimally select the number of control points to be used for the least squares fitting of spline surfaces to a given set of data points. Two criteria were considered - the error in the fitting and the fairness of the surface. It is seen that using 6 control points in the two directions results in a fitted surface patch which is optimal with regards to the surface error and the fairness. The same procedure can also be expanded for use with unequal number of control points in the two parametric directions for the patch. The procedure presented will be useful in the development of a reverse engineering system which generates solid models of digitized objects which incorporates parametric spline surfaces.

REFERENCES

[Farin87] Farin, G., Rein, G., Sapidis, N., Worsey, A.J.: Fairing cubic B-spline curves, Computer aided Geometric Design, Vol. 4, pp. 91-103, 1987

[Farin89] Farin, G., Sapidis, N.: Curvature and the Fairness of Curves and Surfaces, IEEE Computer Graphics and Applications, pp. 52-57, 1989

[Haftk92] Haftka, R.T., Gurdal, Z.: Elements of Structural Optimization, Kluwer Academic Publishers, Boston, USA, 1992.

[Izumi79] Izumida, K., Matida, Y.: Ship hull definition by surface techniques for production use, Computer Applications in the automation of Shipyard Operation and Ship Design III, North Holland Publishing Company, pp. 95-104, 1979.

[Jendo83] Jendo, S., Marks, W: Foundations of Structural Multicriteria Optimization, Proceedings of the IFIP WG 5.2 Working Conference on Optimization in Computer Aided Design, Lyon, France October 24-26, pp. 109-138, 1983.

[Klass80] Klass, R.: Correction of local surface irregularities using reflection lines, Computer Aided Design, Vol.12, No.2, pp.73-77, 1980.

[Lee89] Lee, E.T.Y.: Choosing nodes in Parametric curve Interpolation, Computer Aided Design, Vol.21, No.6, pp. 363-368, 1989.

[Munch78] Munchmeyer, F., Lau, G.K.H.: On the Iterative Design of Smooth Patched Surfaces, IEEE Computer Vision and Image processing, 1978.

[Piegl91] Piegl, L.: On NURBS: A Survey, IEEE Computer Graphics and Applications, Vol. 11, No. 5, pp.55-71, 1991.

[Punta94] Puntambekar, N.V., Jablokow, A.G., Sommer III, H.J.: A Unified Review of 3D Model Generation for Reverse Engineering, Computer Integrated Manufacturing Systems Journal, Vol.7, No. 4, pp. 259-268, 1994.

[Roger90] Rogers, D. F., Adams, J. A.: Mathematical Elements For Computer Graphics, Mcgraw Hill, New York, 1990.