S. K. Boyd
Department of Mechanical Engineering,
email: skboyd@ucalgary.ca

J. L. Ronsky
Department of Mechanical Engineering
email: jlronsky@ucalgary.ca

University of Calgary,
2500 University Drive, N.W.,
Calgary, T2P 1N4 Canada

D. D. Lichti
School of Spatial Sciences,
Curtin University of Technology,
GPO Box U1987,
Perth, WA 6845 Australia

D. Šalkauskas
Department of Mathematics and Statistics

M. A. Chapman
Department of Geomatics

University of Calgary,
2500 University Drive, N.W.,
Calgary, T2P 1N4 Canada

Joint Surface Modeling With Thin-Plate Splines

Mathematical joint surface models based on experimentally determined data points can be used to investigate joint characteristics such as curvature, congruency, cartilage thickness, joint contact areas, as well as to provide geometric information well suited for finite element analysis. Commonly, surface modeling methods are based on B-splines, which involve tensor products. These methods have had success; however, they are limited due to the complex organizational aspect of working with surface patches, and modeling unordered, scattered experimental data points. An alternative method for mathematical joint surface modeling is presented based on the thin-plate spline (TPS). It has the advantage that it does not involve surface patches, and can model scattered data points without experimental data preparation. An analytical surface was developed and modeled with the TPS to quantify its interpolating and smoothing characteristics. Some limitations of the TPS include discontinuity of curvature at exactly the experimental surface data points, and numerical problems dealing with data sets in excess of 2000 points. However, suggestions for overcoming these limitations are presented. Testing the TPS with real experimental data, the patellofemoral joint of a cat was measured with multistation digital photogrammetry and modeled using the TPS to determine cartilage thicknesses and surface curvature. The cartilage thickness distribution ranged between 100 to 550 μm on the patella, and 100 to 300 μm on the femur. It was found that the TPS was an effective tool for modeling joint surfaces because no preparation of the experimental data points was necessary, and the resulting unique function representing the entire surface does not involve surface patches. A detailed algorithm is presented for implementation of the TPS.

Introduction

Mathematical surface models based on experimental measures are important for biomechanical studies because they can be used to describe complicated surface geometry, and to determine parameters such as surface normals and curvature. This is useful, for example, for finite element studies of joint mechanics, for estimating joint contact (Scherrer et al., 1979; Ateshian et al., 1994) or cartilage thickness (Ateshian et al., 1991; Soslowsky et al., 1992), for estimating joint congruency (Ateshian et al., 1992), and for characterizing surfaces based on curvature analysis (Kwak et al., 1997). Surface geometry, normals, and curvature can be calculated at any point on a joint using a mathematical surface model.

Representing a joint surface using a mathematical model based on discrete measurements is challenging because biological surfaces exhibit complicated surface geometry. Typical difficulties encountered with mathematical surface models are discussed by Schut (1976) and Ateshian (1993). These include the difficulty of modeling scattered (nonuniformly distributed) surface measurements, modeling experimental data that are arranged in sparse grids (i.e., data unavailable in rows and columns to complete a rectangular grid), and modeling unordered data (no identification of data points into defined rows and columns). Also, if there is a significant level of measurement error in the surface points, it may be advantageous to incorporate a model that fits (i.e., smooths) the surface data rather than interpolating the data. Discontinuous changes in curvature, which may be the case in joints that have sustained an injury, are problematic for surface models.

Various methods for modeling joint three-dimensional surfaces have been proposed, each having their own strengths and limitations. Scherrer and Hillberry (1979) used piecewise bi-cubic surface patches; however C^2 continuity (continuous second derivatives, implying a smooth surface) was not maintained across patch boundaries. Cubic B-splines exhibiting C^2 continuity were used by Ronsky (1994) to model along MRI slice edges; however, linear interpolation was used in the transverse direction to the slices, and these did not have C^1 continuity (continuous first derivatives). A technique for modeling a surface with quintic B-splines having C^4 continuity in both parametric directions has been proposed by Ateshian (1993). This method has been developed to deal with sparse data, and somewhat nonuniformly distributed data points. However, the primary limitation of this technique, and all the others proposed, is that they are based on tensor products of curve-fitting splines. The limitation of tensor products is that the surface data must be nominally gridded (i.e., not randomly distributed) and arranged in a nonsparse grid, and require administration of the surface patches. Methods have been developed to overcome some of the limitations surrounding sparse grids (e.g., Ateshian, 1995), which require, for example, techniques such as B-spline trimming curves.

For data that are scattered, modeling techniques based on radial basis functions can be successful. These represent the surface in terms of a low degree polynomial surface, to which is added a surface interpolating the difference between this polynomial and the data at the data sites. The latter surface is expressed in terms of translates to the data sites of a suitable radially symmetric function $f(x, y) = \Phi(r)$, $r^2 = x^2 + y^2$, chosen by the user. Examples of such functions are the function $r^2 \ln(r)$ associated with the thin-plate spline and the rotated Gaussian $e^{-\alpha r^2}$, $\alpha > 0$ (α is a shape parameter). If there are n distinct data sites (x_i, y_i), $1 \le i \le n$, then there are n translates $f_i(x, y) = f(x - x_i, y - y_i)$, and the representation of the surface involves a linear combination of them, with coefficients determined by solving a system of linear equations. As will be shown subsequently, this system involves the

Contributed by the Bioengineering Division for publication in the JOURNAL OF BIOMECHANICAL ENGINEERING. Manuscript received by the Bioengineering Division May 28, 1998; revised manuscript received May 18, 1999. Associate Technical Editor: K.-N. An

symmetric matrix whose elements are $a_{ij} = f_i(x_j, y_j)$. This matrix requires certain properties related to invertibility, and consequently the choice of $\Phi(r)$ is not arbitrary. Ultimately the surface is described by a single equation of the form $z = S(x, y)$. The surface does not involve any patches, unlike spline methods based on tensor products.

A classic interpolating function S is the thin-plate spline (TPS) (Meinguet, 1979; Lancaster and Salkauskas, 1986), which is a two-dimensional analogue of the natural cubic spline. The TPS makes use of the radial function $\Phi(r) = r^2 \ln(r)$ mentioned above. It arose originally as a solution of the following idealized mechanical problem: Determine the shape of a thin plate of infinite extent that deforms only in bending (no shear), passes through a number of specified data points, and has minimal linearized bending energy. By this it is meant that the plate has only to undergo small deflections. For a surface of the form $z = f(x, y)$ the energy of the function f is given by the functional

$$E(f) = \int\int_{\mathbb{R}^2} \left\{ \left[\frac{\partial^2 f}{\partial x^2} \right]^2 + 2 \left[\frac{\partial^2 f}{\partial x \partial y} \right]^2 + \left[\frac{\partial^2 f}{\partial y^2} \right]^2 \right\} dx dy, \quad (1)$$

where $\mathbb{R}^2$ is the entire x–y plane.

For the energy $E(f)$ to be finite, it is necessary for the surface $z = f(x, y)$ to flatten out far away from the data sites in the sense of having small curvature. If the data consist of the values z_i at the points (x_i, y_i), $i = 1, \ldots, n$, for n data points, then there is a unique function S such that $z_i = S(x_i, y_i)$, and for which $E(f)$ is minimized, i.e., $E(S) \leq E(f)$. It consists of a first-degree polynomial to which is added a linear combination of the translates to the data sites of the radial function $r^2 \ln(r)$. The derivation of this result is complex and a simplified discussion can be found in Lancaster and Salkauskas (1986).

In conditions where smoothing is desired, it is possible to replace the interpolation conditions $z_i = S(x_i, y_i)$ by a least-squares criterion. This results in a so-called smoothing spline. The idea is to find a function S so that the functional

$$J(f) = E(f) + \sum_{i=1}^{n} w_i [f(x_i, y_i) - z_i]^2 \quad (2)$$

is least when $f = S$. The positive weights, w_i, determine how close to z_i the values $f(x_i, y_i)$ will be when the minimization is carried out. For noisy data z_i, w_i should be inversely proportional to the variance of z_i. For the case where the weights are large, the minimization will focus on the dominant term of $J(f)$ making $f(x_i, y_i)$ close to z_i, and we will obtain a function S much like a thin-plate spline that smooths, rather than interpolates, the data. Alternatively, for the case where the weights are close to zero, there is little regard for the data z_i, and the minimization focuses on the energy term resulting in an almost planar surface will result.

The purpose of this study is to present the thin-plate spline for the mathematical modeling of joints in biomechanical applications in a simple and efficient manner. It is hypothesized that the TPS can provide an accurate mathematical model of a joint surface. The TPS modeling characteristics will be quantified using analytical surface data. The technique will be applied to surface data from a cat patellofemoral (PF) joint from which useful biomechanical information such as cartilage thickness and surface curvature measures will be determined. Detailed methods used for interpolating and smoothing surfaces, as well as determining higher order derivatives used for surface normals and surface curvature, will be presented. It will be shown that the technique is easy to implement for modeling unordered, scattered data points to produce a mathematical surface model that is valuable for biomechanical research.

Methods

Surface Data Sets. An analytical surface was developed to quantify the interpolating and smoothing characteristics of the TPS, as well as the influence of modeling the data in Cartesian versus cylindrical coordinate systems. The surface was designed to have topological features similar to a natural joint. A surface of revolution about the x axis (± 80 deg) was created based on a fourth-order polynomial

$$z = -(\tfrac{1}{6}x)^4 + 3(\tfrac{1}{6}x)^2 + 6, -10 \leq x \leq 10 \text{ (units of mm)}. \quad (3)$$

Five sets of randomly distributed data points (x_i, y_i, z_i) that lay on the analytical surface were created ranging from 200 to 2000 points to test the error of the TPS interpolating surface model. Similarly, for TPS smoothing tests, two sets of 1000 randomly distributed data points were generated with added random (white) noise in the z direction of 50 μm and 100 μm standard deviation (S.D.). Included in the random points on the analytical surface were points added to the edge to ensure the entire surface was represented.

Illustrating the practical use of the TPS, experimental surface data were collected from the PF joint of a cat, including both the subchondral bone and cartilage surfaces, using the technique of multistation digital photogrammetry (MDPG) (Ronsky et al., 1999). The surface data were measured in a Cartesian coordinate system. The MDPG technique provides a three-dimensional coordinate error estimate for each surface point, as well as an averaged coordinate error estimate for all the points on each surface. These error estimates were utilized for the weighted least-squares smoothing implementation of the TPS.

Thin-Plate Spline Model. The description of the TPS model is provided in detail so that the interested reader may write software to implement the TPS.

It can be shown (Meinguet, 1979) that the thin plate spline surface has an equation of the form

$$S(x, y) = \sum_{i=1}^{n} c_i f_i(x_i, y_i) + c_{n+1} + c_{n+2}x + c_{n+3}y \quad (4)$$

where the c_i are certain constants and the functions f_i are translates of $r^2 \ln(r)$ to the data sites (x_i, y_i), $i = 1, \ldots, n$, for n data points. Thus,

$$\Phi(r) = f_i(x, y) = r_i^2 \ln(r_i), \quad \text{with } r_i^2 = (x - x_i)^2 + (y - y_i)^2. \quad (5)$$

The f_i are indeterminate at (x_i, y_i); an application of l'Hôpital's rule shows that the value should be defined to be zero. This makes the f_i continuous. Also, the first partial derivatives are continuous, but the second derivatives are discontinuous at the data sites.

The $n + 3$ coefficients in S are only partially determined by the n interpolation conditions

$$z_i = S(x_i, y_i), \quad i = 1, \ldots, n. \quad (6)$$

An additional *polynomial precision* condition is that if the points (x_i, y_i, z_i) lie on a plane, then $S(x, y)$ reduces to the equation of that plane by virtue of the vanishing of the coefficients c_i, $i = 1, \ldots, n$. It is convenient to use matrix notation to describe $S(x, y)$ and the method for determining its coefficients:

$$S(x, y) = \mathbf{f}\mathbf{c}^T, \quad (7)$$

where $\mathbf{f} = [f_1(x, y), \ldots, f_n(x, y), 1, x, y]$ and $\mathbf{c} = [c_1, \ldots, c_{n+3}]$.

The interpolation and polynomial precision conditions imply that the coefficient vector $\mathbf{c}$ is a solution of the following $(n + 3) \times (n + 3)$ linear system, conveniently written in block form as

$$\begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix}. \quad (8)$$

The entries in A are $a_{ij} = f_i(x_j, y_j)$, $i \neq j$; $a_{ii} = 0$; $i, j = 1, \ldots, n$. This matrix is in fact symmetric. As mentioned earlier, A, which is created from the radial function Φ, has to be such that the matrix in (8) is invertible. The other vectors and matrices have the forms

$$B^T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ y_1 & y_2 & y_3 & \cdots & y_n \end{bmatrix}, \quad \mathbf{c}_1^T = [c_1, c_2, \ldots, c_n],$$

$$\mathbf{c}_2^T = [c_{n+1}, c_{n+2}, c_{n+3}], \quad \mathbf{z} = [z_1, z_2, \ldots, z_n]^T. \tag{9}$$

The first "block" equation $A\mathbf{c}_1 + B\mathbf{c}_2 = \mathbf{z}$ represents the interpolation conditions, and the second one, $B^T\mathbf{c}_1 = \mathbf{0}$, is the polynomial precision condition. The coefficients in $\mathbf{c}$ can be determined from this system by using any linear equation solver. These solvers use techniques that may be more efficient than matrix inversion. Formally, the solution may be written

$$\mathbf{c} = \begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix}. \tag{10}$$

If smoothing is desired, then the functional $J(f)$ can be minimized. The optimal function is closely related to the thin-plate spline (Wahba, 1990). It has the form $S(x, y) = \mathbf{fc}^T$, where $\mathbf{f}$ is as in Eq. (7), and the system (8) is modified to

$$\begin{bmatrix} A + D & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix}, \tag{11}$$

where D is a diagonal matrix containing the reciprocals of the weights w_i of Eq. (2),

$$D = \text{diag}\,(1/w_1, \ldots, 1/w_n). \tag{12}$$

The weights w_i are a function of the standard deviation of the data, σ_i^2, and a smoothing parameter, λ, expressed as

$$w_i = \lambda/\sigma_i^2. \tag{13}$$

Derivatives. The determination of surface normals and curvature requires the first and second partial derivatives of the surface function S in Eq. (7). The coefficient vector $\mathbf{c}$ of Eq. (7) has already been obtained by solving the system in Eq. (8) or (11); therefore only partial derivatives of the functions f_i in the vector $\mathbf{f}$ need to be calculated. The first partial derivatives are calculated as follows:

$$\frac{\partial}{\partial x} S(x, y) = [f_1^x \quad f_2^x \quad f_3^x \quad \cdots \quad f_n^x \quad 0 \quad 1 \quad 0]\mathbf{c}^T,$$

$$f_i^x = \frac{\partial f}{\partial x} = (x - x_i)\ln(r_i^2) + (x - x_i) \tag{14}$$

A similar calculation yields $\partial S/\partial y$ and f_i^y.

The second partial derivatives are:

$$\frac{\partial^2}{\partial x^2} S(x, y) = [f_1^{xx} \quad f_2^{xx} \quad f_3^{xx} \quad \cdots \quad f_n^{xx} \quad 0 \quad 0 \quad 0]\mathbf{c}^T,$$

$$f_i^{xx} = \frac{\partial^2 f}{\partial x^2} = \frac{2(x - x_i)^2}{r_i^2} + \ln(r_i^2) + 1, \tag{15}$$

and similarly for $\partial^2 S/\partial y^2$ and f_i^{yy}. Finally,

$$\frac{\partial^2}{\partial x \partial y} S(x, y) = [f_1^{xy} \quad f_2^{xy} \quad f_3^{xy} \quad \cdots \quad f_n^{xy} \quad 0 \quad 0 \quad 0]\mathbf{c}^T,$$

$$f_i^{xy} = \frac{\partial^2 f}{\partial x \partial y} = \frac{2(x - x_i)(y - y_i)}{r_i^2}. \tag{16}$$

Analytical Surface Modeling. Quantitative analysis of the TPS interpolating and smoothing characteristics was done using the analytical surface data. For each of the five randomly distributed data set ranging in size from 200 to 2000 points, the interpolating TPS was used to model the data in either the Cartesian or cylindrical coordinate systems. The choice for the cylindrical coordinate system was determined by finding the axis of a cylinder that best fits the data using least-squares estimation. Comparison of the TPS surface to the analytical surface was done by resampling in a grid pattern from the TPS surface and estimating the error at each resampling point relative to the true analytical surface.

The smoothing characteristics of the TPS were tested using the analytical data sets of 1000 points with added random noise. Selection of the smoothing parameter, λ, involved systematically applying a range of values, and measuring the fitting error between the smoothed TPS surface and the noisy data. With *a priori* knowledge of the noise in the data (50 μm S.D. or 100 μm S.D.) it is possible to interpolate an appropriate smoothing parameter, λ. This is a standard method for determining the smoothing factor for experimental data based on *a priori* estimated noise (e.g., Ateshian, 1993). We also investigated the proximity of the smoothed TPS surface (based on the noisy data) to the true analytical surface for the same range of smoothing parameters, λ. The percent error was calculated between an optimally smoothed surface, which occurs presumably when the TPS represents the true analytical surface with minimum error, and the smoothed surface based on the smoothing factor determined from the *a priori* known noise.

Curvature maps of the interpolated and smoothed noisy data were used to provide information regarding the smoothness of the TPS surface that is difficult to detect on the plotted surfaces alone. Gaussian curvature, which is based on the product of the principal curvatures ($K = \kappa_1\kappa_2$), is sensitive to discontinuities (either natural or caused by measurement errors) in the surface curvature. The principal curvatures (κ) and their orientation (h) were determined by solving for the roots of the following two equations (Mortenson, 1985) at every re-sampled point on the surface based on partial derivatives determined from the TPS calculations:

$$(EG - F^2)\kappa^2 - (EN + GL - 2FM)\kappa + (LN - M^2) = 0,$$

$$(FN - GM)h^2 - (EN - GL)h + (EM - FL) = 0. \tag{17}$$

Here,

$$E = \mathbf{S}^x \cdot \mathbf{S}^x, \quad F = \mathbf{S}^x \cdot \mathbf{S}^y, \quad G = \mathbf{S}^y \cdot \mathbf{S}^y,$$

$$L = \mathbf{S}^{xx} \cdot \mathbf{n}, \quad M = \mathbf{S}^{xy} \cdot \mathbf{n}, \quad N = \mathbf{S}^{yy} \cdot \mathbf{n},$$

and,

$$\mathbf{S}^x = \begin{bmatrix} 1 & 0 & \dfrac{\partial S}{\partial x} \end{bmatrix}, \quad \mathbf{S}^y = \begin{bmatrix} 0 & 1 & \dfrac{\partial S}{\partial y} \end{bmatrix},$$

$$\mathbf{S}^{xx} = \begin{bmatrix} 0 & 0 & \dfrac{\partial^2 S}{\partial x^2} \end{bmatrix}, \quad \mathbf{S}^{yy} = \begin{bmatrix} 0 & 0 & \dfrac{\partial^2 S}{\partial y^2} \end{bmatrix},$$

$$\mathbf{S}^{xy} = \begin{bmatrix} 0 & 0 & \dfrac{\partial^2 S}{\partial x \partial y} \end{bmatrix},$$

where $\mathbf{n}$ is a surface normal vector, and "$\cdot$" indicates a dot product of two vectors.

Experimental Surface Modeling. The interpolating TPS was used to model the data from the cat PF joint data in the Cartesian coordinate system. The modeled surfaces were resampled in a grid pattern with spacing of 0.5 mm. Because the modeled surface extends beyond the input surface data to infinity, an automatic method of discarding resampled data points far from the original data points was devised. For every resampled point, an algorithm (Matlab, v5, Mathworks, Inc., Natick, MA) calculated the distance to the three closest original data sites. If the average distance to those three points was greater than the mean distance between the entire set of points, the resampled point was discarded.

Cartilage thickness was determined for the cat PF joint, including the patella and the femur condylar groove. The thickness measurement was made using a computer algorithm in which thickness was defined as the distance along the normal, **n**, projected from a resampled point on the bone surface to its intersection with the cartilage surface. The resampled points on the two TPS surfaces do not need to be matched. When **n** is normalized ($\mathbf{n}/\|\mathbf{n}\|$), then the distance, t, is measured in the same units as the surface measurements (i.e., mm). The distance t at which the normal vector from the bone intersects with the cartilage surface is a direct measure of the thickness at one point on a surface (Fig. 1). Taking several such estimations yields an overall map of the cartilage thickness (Ateshian et al., 1991).

Solving for the cartilage thickness is possible with the TPS model on a point-by-point basis by determining the distance t for every resampled point on the bone. First, the line normal to the bone surface is parametrized. If (x_o, y_o, z_o) is a resampled point on the bone surface, and (x_l, y_l, z_l) is a point a distance t along the surface normal line, **n**, then the parametrized line is

$$x_l = \mathbf{n}_x t + x_o, \quad y_l = \mathbf{n}_y t + y_o, \quad z_l = \mathbf{n}_z t + z_o. \tag{18}$$

Because the surface heights, z, are represented by a function of two variables (x, y), a closed equation can express the intersection of the line with the surface. Intersection occurs when the height, z, of the cartilage surface, S_c, at the x_l and y_l values along the normal line equals the z_l value of that line:

$$S_c(x_l, y_l) - z_l = 0, \tag{19}$$

or, after substitution of the parametrized line in terms of t,

$$S_c(\mathbf{n}_x t + x_o, \mathbf{n}_y t + y_o) - (\mathbf{n}_z t + z_o) = 0.$$

Solving for the only unknown, t, in the proximity equation is done using Newton's method after re-arranging the equation:

$$t' = t - \frac{S_c(t) - (\mathbf{n}_z t + z_o)}{\frac{d}{dt}\left(S_c(t) - (\mathbf{n}_z t + z_o)\right)} \tag{20}$$

where, using the chain rule,

$$\frac{dS_c}{dt} = \left(\frac{\partial S_c}{\partial x}\right)\left(\frac{\partial x}{\partial t}\right) + \left(\frac{\partial S_c}{\partial y}\right)\left(\frac{\partial y}{\partial t}\right).$$

For each discrete resampled point on the bone surface, a measure of cartilage thickness was estimated.

Results

The TPS was successful at creating a mathematical model of all the analytical and surface data sets. Test results are presented for the TPS with the randomly distributed data sets (200 to 2000 points) lying on the analytical surface (Fig. 2(*a*)) to determine the TPS modeling error relative to the true analytical surface (Table 1). The mean error decreased as the number of points in the data set increased. This result confirms what is expected of all reasonable mathematical models for complex surfaces, that increasing the number of data points improves the representation of the true analytical surface. An interesting result was that the error was decreased by an order of magnitude when the data sets were modeled in a cylindrical coordinate system. Regionally, in the Cartesian coordinate system, the largest errors occurred near the edges of the analytical surface when the tangent of the surface exceeds 75 deg relative to the X–Y plane (Fig. 2(*b*)); however, in the cylindrical coordinate system these errors were significantly reduced (Fig. 2(*c*)). The Gaussian curvature map (Fig. 2(*d*)) of the surface represented in the Cartesian coordinate system illustrates that the TPS surface is smooth, with the exception of the regions near the edge where fluctuations in curvature were present.

Selection of an appropriate smoothing factor for the analytical data with added noise was investigated with the fitting-error curves in Fig. 3(*a*). As expected, increasing the smoothing factor toward infinity creates a surface that more closely interpolates the noisy data (i.e., solid lines approach zero). Also, the error with respect to the true analytical surface (dashed lines) reaches a minimum and then increases asympotically to the noise level of the input data (as λ approaches infinity). Basing the selection of the smoothing factor on the *a priori* known error in the analytical data (in this case, 50 μm S.D.) produces a smoothing factor ($\lambda = 1.37$) with surface modeling errors within 3 percent of the optimal error, and the surface is slightly oversmoothed. Optimal smoothing for this analytical data occurs when the curves representing the proximity of the smoothed TPS surface to the true analytical surface (dashed lines) are minimized. The Gaussian curvature when the noisy surface data are smoothed at $\lambda = 1.37$ has reduced fluctuations (Fig. 3(*b*)) compared to the unsmoothed surface (Fig. 3(*c*)).

A summary of the experimental data sets obtained using MDPG is presented in Table 2. Plots of the original surface data points and the TPS surfaces are presented for the cat femur and patella subchondral bone (Fig. 4). The cartilage thickness of the patella and femur in the cat PF joint was determined. The cartilage thickness pattern represented with gray-scale contours is markedly different for the femur (Fig. 5) and patella (Fig. 6). Cartilage thickness ranges from between 100 to 550 μm for the patella, and 100 to 300 μm for the femur.

Gaussian surface curvature maps (similar to Fig. 2(*d*)) were calculated for all four of the experimental data surface sets (Table 2), which revealed that when no smoothing was applied (i.e., interpolation, or $\lambda = \infty$) then there were curvature fluctuations near some of the data points. However, when the surfaces were smoothed, these fluctuations decreased. The selection of the smoothing factor was based on *a priori* knowledge from the MDPG technique measurement error. For example, the MDPG error was 10.6 μm for the patellar subchondral bone surface (Table 2), and the same technique as shown in Fig. 3(*a*) was used to select a smoothing parameter of $\lambda = 2.8$.

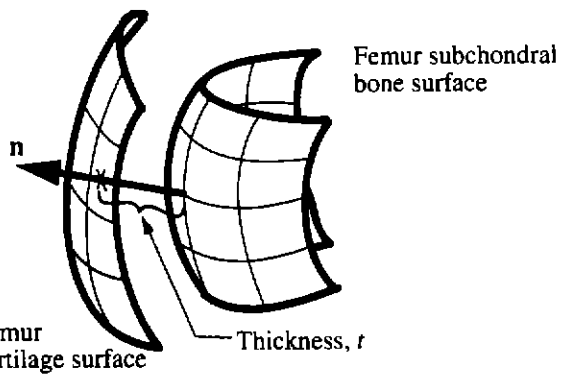
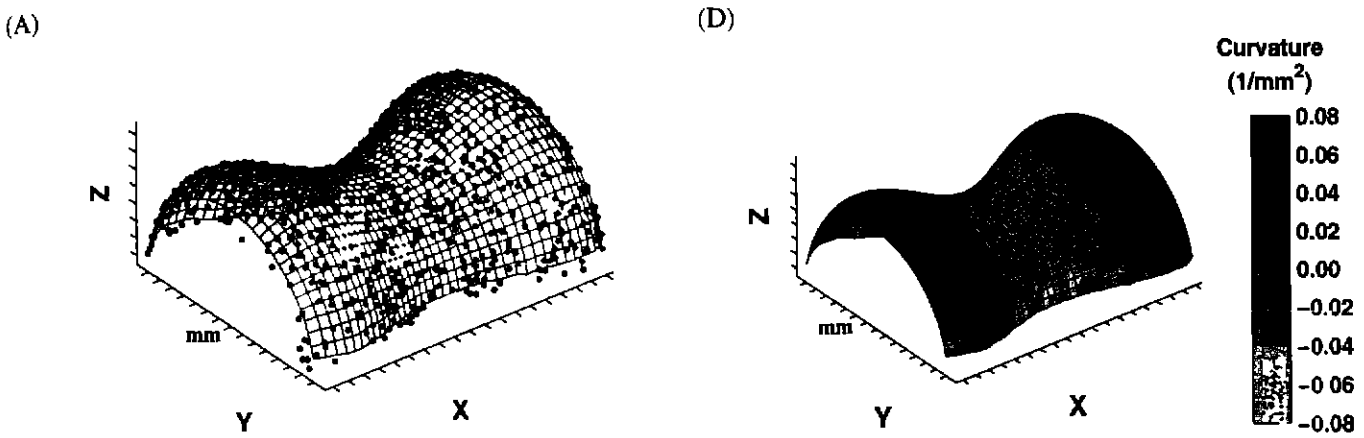


Fig. 1 The cartilage thickness is calculated by determining the intersection distance, t, along the normal, n, from each resampled point on the bone surface to the cartilage surface. Alternatively, the same method may be used to find the proximity of one surface to another to estimate joint contact.

Discussion

The radial basis function method is a new alternative in biomechanical research for joint surface modeling that can interpolate surface data using the TPS, and can smooth surface data by incorporating least-squares smoothing. Using random, unordered analytical data, it was shown that the TPS represents the true analytical surface with low error. Practical use of the TPS on the experimental data collected with MDPG of the cat PF joint was demonstrated, and important joint surface characteristics such as cartilage thickness and surface curvature were determined.

The ability of the radial basis function model to create a mathematical surface model $z = S(x, y)$ from unordered, scattered data (Fig. 2(*a*)) simplifies the modeling process. The TPS method does



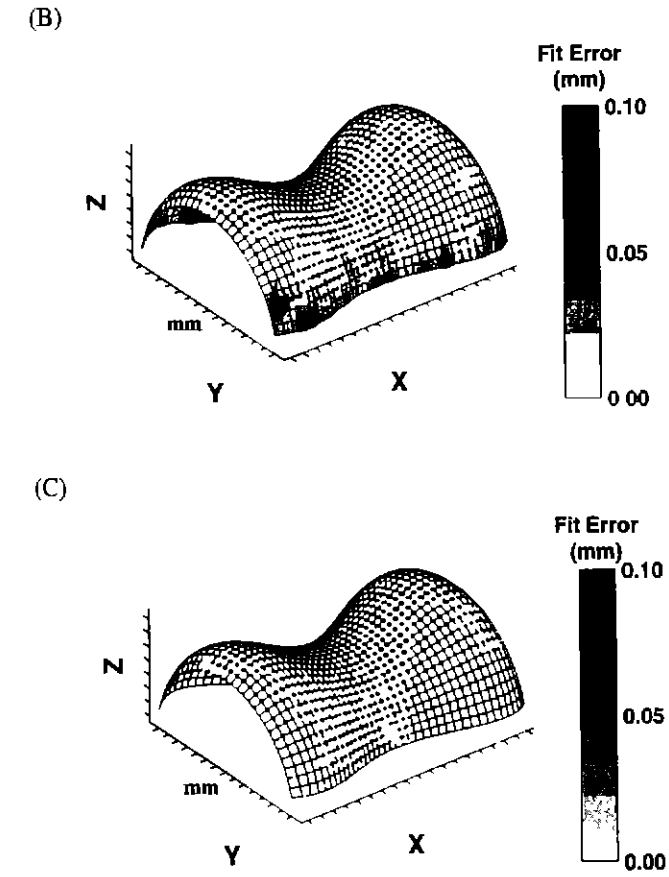


Fig. 2 (*A*) **Randomly distributed data points that lie on an analytical surface (1000 points in this example) and the TPS mesh generated from modeling these points.** (*B*) **The random data points are modeled in the Cartesian coordinate system, and errors between the TPS mesh and the true analytical surface are represented as gray-scale contours on the mesh.** (*C*) **The same points are modeled in the cylindrical coordinate system with error represented as gray-scale contours, which are negligible.** (*D*) **The Gaussian curvature for the data modeled in the Cartesian coordinate system. The surface is smooth except at the edges.**

not require specification of control points, knot sequences, and assignment of parametric coordinates to experimental data points as with tensor product based methods. This is useful, for example, to avoid a frequently encountered difficulty with experimental data from photogrammetric measurements (e.g., Ateshian et al., 1991; Boyd et al., 1997) where multiple grid sets from different regions of the same joint overlap (i.e., Fig. 4(*b*)). With the TPS technique, no organization of data points is necessary before a full mathematical description of the surface can be obtained and represented with a single function, $z = S(x, y)$. The use of a single function representing the entire joint surface without piecewise representation simplifies analysis compared to tensor product methods because the organizational aspect of analyzing surfaces on a patch-by-patch basis is eliminated.

Modeling problems may occur if two adjacent data points representing a smooth articular surface have different heights, causing a surface discontinuity. Unless a physical discontinuity is present, the difference in height is likely due to error in the experimental measurement technique. Applying a smoothing factor with the TPS becomes necessary so that a smooth mathematical representation of the joint is possible. The ability to smooth raw surface data is an important feature of all surface models because often there is measurement error inherent in the data. Selection of an appropriate smoothing factor controls the amount of fitting error between the mathematical surface and raw surface data. Potential extremes to the selection of the smoothing factor include a factor of infinity ($\lambda = \infty$), which results in no smoothing (raw data is interpolated), and a factor near zero ($\lambda \approx 0$), which results in oversmoothing and the mathematical model degenerates into a plane. Selection of an appropriate smoothing factor was tested on the analytical data, which involved comparison of the fitting error to the *a priori* known noise in the data (Fig. 3(*a*)). It appears that this selection criterion results in slight oversmoothing of the surface, and that the optimal smoothing factor is reached when the TPS surface proximity is minimized (at a larger value of λ). The error due to nonoptimal selection of the smoothing factor based on the analytical data set was less than 3 percent of the minimum possible error for data with 50 μm S.D. of noise. The potential errors, however, increase with increasingly noisy data, and the error was as high as 15 percent for data with 100 μm S.D. of noise. However, with experimental data, rather than analytical data, there is no knowledge of the proximity of the surface model to the true surface. Therefore, estimation of the smoothing factor based on known (or estimated) measurement error is a useful, objective method. A simple subjective check of the smoothing factor can be done by inspecting the resulting curvature of the smoothed surface.

The TPS model smooths noisy surface data by combining the minimal bending energy condition and a least-squares fitting error. Unlike polynomial smoothing in the sense of regression, the smoothing spline involves as many basis functions (i.e., $n + 3$) as the interpolating spline. Polynomial interpolants and least-squares approximants of high degree are prone to large oscillations and are

Table 1 The error of the TPS surface based on randomly distributed data sets lying on a known analytical surface with sizes ranging between 200 and 2000 points (mean, standard deviation). The identical data points were modeled in both the Cartesian and cylindrical coordinate systems.

Coordinate system	Fit Error mm	200 Pts	500 Pts	1000 Pts	1500 Pts	2000 Pts
Cartesian	Mean (S.D.)	0.0057 (0.0241)	0.0027 (0.0084)	0.0093 (0.0223)	0.0056 (0.0156)	0.0041 (0.0120)
Polar	Mean (S.D.)	0.0028 (0.0054)	0.0008 (0.0027)	0.0003 (0.0003)	0.0003 (0.0002)	0.0003 (0.0002)

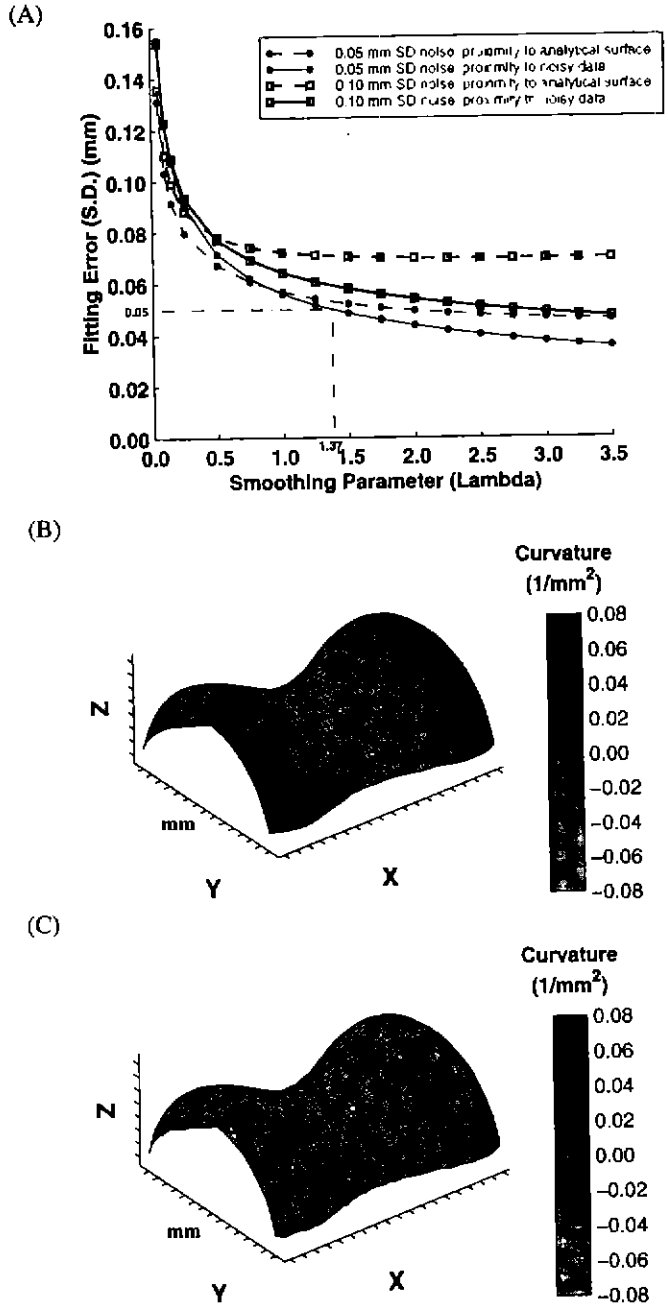


Fig. 3 (*A*) The fitting error of the weighted TPS smoothed surface models based on data with added noise (50 μm S.D.—solid circles, and 100 μm S.D.—hollow squares) while varying the smoothing parameter (0 to 3.5, corresponding to decreasing smoothing). The solid lines represent the proximity of the surface to the noisy data, and the dashed lines represent the proximity of the surface to the underlying analytical surface without added noise. A smoothing factor of 1.37 is chosen for the data with 50 μm S.D. added noise based on the corresponding curve. (*B*) The gaussian curvature is calculated for the TPS surface while applying a smoothing factor of 1.37, and (*C*) the same surface without smoothing (interpolating) is presented to show how the surface is rippled compared to the smoothed surface (*B*).

not recommended. The minimization of the functional J helps to reduce oscillations in the fitted surface, although these cannot be avoided entirely.

Furthermore, with regards to the concept of local support, both the TPS and cubic (or quintic) splines share the property that the perturbation of the data at one site propagates throughout the domain, although it attenuates rapidly. Alternatively, univariate splines have a compactly supported basis for their representation utilizing B-splines. The compact support makes the computation more efficient than the TPS, which involves full, rather than sparse, matrices. However, with widely available efficient computers, this difference in processing speed has little practical influence.

Due to the rapidly attenuated TPS, there is only a limited region of practical local support. This region can be estimated with a simple test, which involves modeling an impulse function. This is where data representing a planar surface include a single point that lies one unit out of the plane, and the region of the modeled plane that is no longer flat is the region affected by local support. The size of this region will depend on factors such as the point density of the data set and the magnitude of the impulse applied to the point. A quick test with a planar data set having a point distribution similar to the experimental data (Table 2), and an impulse function of magnitude of 1 mm, yielded a radial region of influence approximately 2 mm. Similar tests of local support may be useful with other types of experimental data sets.

Caution must be exercised when modeling some types of surface data sets. First, numerical errors will occur if fewer than three noncollinear points are modeled. Second, numerical difficulties may occur if the data set is very large. This is because radial functions require the solution of a linear system with a full matrix, and round-off errors may occur (Franke and Salkauskas, 1996). The maximum size of the data set is machine and software specific; however, surface data sets with up to 2000 data points were successfully modeled in this study (SGI Indigo 2, Matlab v5). Most surface measurement techniques used in biomechanics result in surface data sets with fewer than 1000 points (Seedhom et al., 1972; Scherrer and Hillberry, 1979; Wismans et al., 1980; Shiba et al., 1988; Wijk, 1980; Ghosh, 1983; Huiskes et al., 1985; Ronsky, 1994), which do not present numerical problems for the TPS technique. In cases of exceptionally large data sets, for example, from laser scanned images, techniques for improving the conditioning of the system may be necessary (Narcowich and Ward, 1992). Additionally, there are some recent developments in the area of positive definite functions with compact support (Wendland, 1995). The use of such functions leads to a sparse matrix to which one may apply special numerical methods. A technique that has been described as "convolution windowing" (Franke and Salkauskas, 1996) may also be used where large data sets are subdivided into regions, or windows, of overlapping data sets.

A potential limitation of the TPS is that although it has infinitely continuous derivatives of all orders over the majority of the surface, it is not C^2 at just the original data points due to discontinuous second derivatives at these points. This may contribute to the curvature fluctuations near some of the original data points in the interpolated experimental data; however, these fluctuations are more likely to due noise in the data. In Fig. 2(*d*), for example, the only fluctuations in curvature occur near the edge of the surface where normal problems associated with edge effects are likely the cause. The lack of C^2 continuity at exactly the data points, therefore, appears to have little effect on the smoothness of the analytical surface. Interestingly, with tensor product methods a similar limitation exists where, although C^2 continuity and higher may exist at knot points and within patches, the continuity is reduced along the intersection lines between adjacent patches.

However, with the TPS, if maintaining C^2 continuity is required even at exactly the data points, then instead of $\Phi(r) = r^2 \ln(r)$,

Table 2 Experimental surface data and corresponding measurement error from MDPG measurements of the four surfaces on the cat patellofemoral joint

		# Pts	Mean S.D. (μm) of measurement error		
			$S.D._x$	$S.D._y$	$S.D._z$
Bone	Patella	75	7.5	6.8	10.6
	Femur	291	8.0	8.7	10.3
Cartilage	Patella	68	11.1	10.0	15.2
	Femur	217	12.7	15.3	17.9

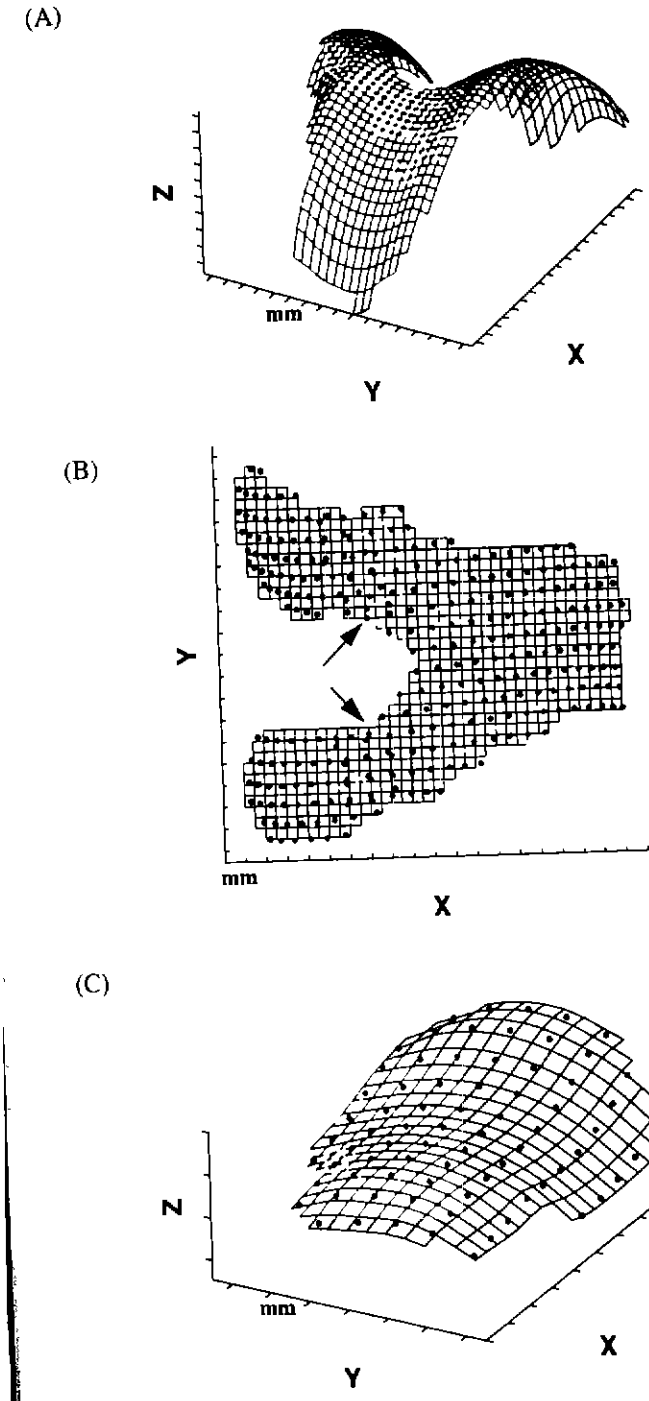


Fig. 4 A TPS mathematical surface model from MDPG measurements (resampled in a grid pattern) of (*A*) the cat distal femur, and (*B*) a top view of the same mesh with the original data points plotted over the resampled mesh. Note that the MDPG measurements were collected from two grid projections on the same surface and result in scattered data at the intersection of the two grid projections indicated by arrows. The TPS mesh with original MDPG measurements of the cat patella (*C*) was collected in one MDPG measurement session.

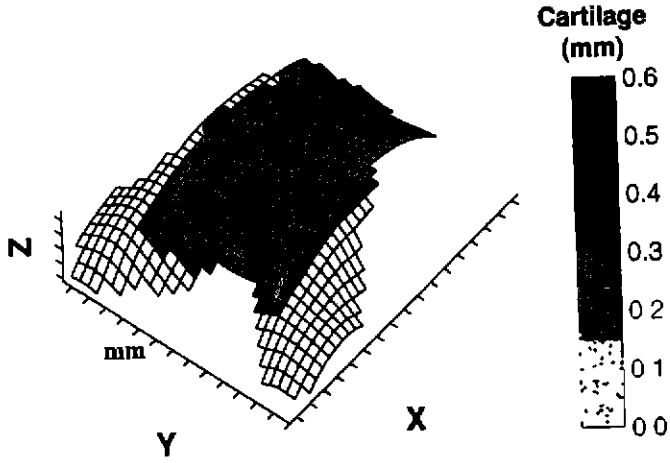


Fig. 5 Cartilage thickness of the cat distal femur represented with gray-scale contours projected onto a surface grid mesh. Thickness measurements were not available in regions where no gray-scale contour is presented.

It is not essential to involve a functional like E or J in the creation of a fitted surface. The choices for $\Phi(r)$, however, are not arbitrary in that the resulting matrices must be invertible. Admissible functions $\Phi(r)$ involve the concepts of positive definiteness and conditional positive definiteness. These ideas are connected to the probabilistic approach found in the method known as Kriging and elsewhere (Wahba, 1990). Examples of additional choices for $\Phi(r)$ are presented for completeness, but have not be utilized in this study:

(*a*) $\Phi(r) = e^{-\alpha r^2}$, $\alpha > 0$, continuity C^∞, polynomial precision condition is not required.

(*b*) $\Phi(r) = r^3$, continuity C^2, polynomial precision condition of at least 1 is required.

The TPS model using $\Phi(r) = r^2 \ln(r)$ was chosen for this study because the minimization of the surface bending is an appealing concept for governing how natural joint surfaces may be shaped. Minimized surface bending would likely result in contact conditions with minimized contact stress; however, testing of this concept would require further investigation.

The analytical data were modeled in both the Cartesian and cylindrical coordinate systems. In the Cartesian coordinate system, TPS errors became significant when the tangent of the modeled surface became too steep; however, the error was reduced by an order of magnitude when the same data were modeled in the cylindrical coordinate system. This occurs mainly because the analytical data is well suited for modeling in a cylindrical coordi-

one may use other radial functions to generate an interpolating or smoothing scheme. Meinguet (1979) presents the radial basis functions associated with functionals J similar in form to Eq. (1), but involving derivatives of orders $m = 2, 3, \ldots$. The resulting surfaces no longer have the minimum bending energy property when $m > 2$, but gain differentiability. Their polynomial precision is $m - 1$, and Eqs. (8) and (11) apply with a change in the entries of A and B^T. The matrix B^T involves the basis functions for bi-variate polynomials of total degree $m - 1$. It is shown that $\Phi(r) = r^{2m-2} \ln(r)$, and there is a discontinuity in the $2m$th derivative.

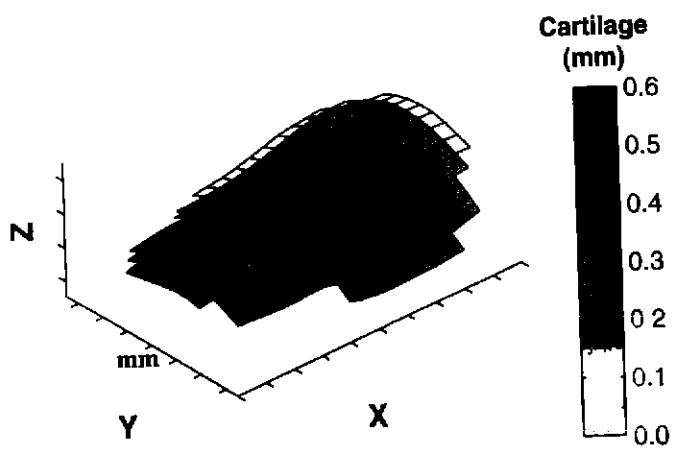


Fig. 6 Cartilage thickness of the cat patella represented with gray-scale contours projected onto a surface grid mesh. Thickness measurements were not available in regions where no gray-scale contour is presented.

nate system. It does illustrate, however, that cylindrical or spherical coordinates may be more suitable if surfaces are highly curved, or if multiple z values exist for one x–y coordinate pair (i.e., highly curved surfaces resembling a sphere). Parametrization of experimental data is also possible with the TPS, which requires assigning experimental data (x, y, z) to parameters (u, v) in parameter space. It is important to note, however, that minimization of the bending energy of the TPS occurs in the coordinate system chosen. For example, in the cylindrical coordinate system (r, θ, z), the surface has an equation of the form $r = C(\theta, z)$, and the minimization of the energy is over the θ–z coordinate plane. Thus, the energy is not minimized in the original (x, y, z) system. Therefore, when surfaces are modeled in coordinate systems other than the Cartesian coordinate system, the concept of minimized bending energy may have little physical relevance. The resulting surface, however, remains smooth as is the case with the analytical data modeled in the cylindrical coordinate system (Fig. 2(*c*)).

Modeling of the experimental PF joint data was done in the Cartesian coordinate system using the TPS without smoothing. This was done because the surface measurements of the subchondral bone and cartilage in the cat PF joint were obtained using MDPG, and had small associated errors (10 to 20 μm) relative to the thickness of the cartilage (100 to 550 μm). The cartilage thickness estimations expressed as contours for the femur (Fig. 5) and patella (Fig. 6) exhibit little fluctuation, which supports the conclusion that smoothing the raw data was not necessary.

The cartilage on the femur was a relatively constant thickness over the entire surface and was in the range of 100 to 300 μm, whereas the cartilage on the patella varied significantly more over the joint surface (100 to 550 μm). In general, the cartilage was thickest slightly proximal to the midpoint of the patella surface (up to 550 μm) and thinnest near the proximal and distal ends of the surface (~100 μm). There are no previous reports of cartilage thickness maps of the cat PF joint in the literature. Comparison of the distribution of the cartilage thickness on the cat patella with studies of the human patella (Ateshian et al., 1991; Eckstein et al., 1996) show similarities where the cartilage is thickest near the midportion of the articulating surface. The distribution of the cartilage thickness on the patella in the dog (Kwak et al., 1993) is different from the cat because the thickest region is located more distally.

In addition to the TPS being practical for estimating entities such as cartilage thickness or curvature in this study, it could also be used for proximity contact measurements (Scherrer et al., 1979; Soslowsky et al., 1992) between opposing surfaces with TPS mathematical models of joint surfaces (Boyd, 1997). Although there are limitations of the TPS, its strength lies in its simplicity for creating a mathematical surface model described as a single function, $z = S(x, y)$, from unordered, noisy, and scattered experimental surface data. The algorithm is simple to implement (based on detailed calculations presented), can be fully automated, and offers a powerful and efficient method for determining surface characteristics of complex joint surfaces.

Acknowledgments

Supported financially by Natural Science and Engineering Research Council of Canada and the Alberta Heritage Foundation for Medical Research. Significant technical assistance was provided by C. Sutherland, D. Young, B. Chaplin, B. Kralovic, and A. Wuilbaut.

References

Ateshian, G. A., 1993, "A B-spline least-squares surface-fitting method for articular surfaces of diarthrodial joints," ASME JOURNAL OF BIOMECHANICAL ENGINEERING, Vol. 115, pp. 366–373.

Ateshian, G. A., 1995, "Generating trimmed B-spline models of articular cartilage layers from unordered 3D surface data points," *Proc. ASME Bioengineering Conference*, pp. 217–218.

Ateshian, G. A., Soslowsky, L. J., and Mow, V. C., 1991, "Quantitation of articular surface topology and cartilage thickness in knee joints using stereophotogrammetry," *J. Biomechanics*, Vol. 24, pp. 761–776.

Ateshian, G. A., Rosenwasser, M. P., and Mow, V. C., 1992, "Curvature characteristics and congruence of the thumb carpometacarpal joint: differences between female and male joints," *J. Biomechanics*, Vol. 25, pp. 591–607.

Ateshian, G. A., Kwak, S. D., Soslowsky, L. J., and Mow, V. C., 1994, "A stereophotogrammetric method for determining *in situ* contact areas in diarthrodial joints, and a comparison with other methods," *J. Biomechanics*, Vol. 27, No. 1, pp. 111–124.

Boyd, S. K., 1997, "A 3D in-situ model for patellofemoral joint contact analysis in the normal and anterior cruciate ligament deficient knee," MSc. Thesis, University of Calgary Press, Calgary, AB.

Boyd, S. K., Ronsky, J. L., Lichti, D. D., Salkauskas, K., and Chapman, M. A., 1997, "Quantification of articular cartilage thickness of the cat patellofemoral joint with multi-station digital photogrammetry and thin-plate spline surface interpolation," *Proc. ASME Bioengineering Conference*, Vol. 35, pp. 43–44.

Eckstein, F., Gavazzeni, A., Sittek, H., Haubner, M., Lösch, A., Milz, S., Englmeier, K. H., Schulte, E., Putz, R., and Reiser, M., 1996, "Determination of knee joint cartilage thickness using three-dimensional magnetic resonance chondrocrassometry," *Magnetic Resonance in Medicine*, Vol. 36, pp. 256–265.

Franke, R., and Salkauskas, K., 1996, "Localization of multivariate interpolation and smoothing methods," *J. Computational and Applied Mathematics*, Vol. 73, pp. 79–94.

Ghosh, S. K., 1983, "A close-range photogrammetric system for 3-D measurements and perspective diagramming in biomechanics," *J. Biomechanics*, Vol. 16, pp. 667–674.

Huiskes, R., Kremers, J., de Lange, A., Woltring, H. J., Selvik, G., and van Rens, J. G., 1985, "Analytical stereophotogrammetric determination of three-dimensional knee-joint geometry," *J. Biomechanics*, Vol. 18, No. 8, pp. 559–570.

Kwak, S. D., Newton, P. M., Setton, L. A., and Grelsamer, R. P., 1993, "Cartilage thickness and contact area determination in the canine knee joint," *Trans. Orthop. Res. Soc.*, Vol. 18, No. 2, p. 351.

Kwak, S. D., Colman, W. W., Ateshian, G. A., Grelsamer, R. P., Henry, J. H., and Mow, V. C., 1997, "Anatomy of the human patellofemoral joint articular cartilage: surface curvature analysis," *J. Orthop. Res.*, Vol. 15, pp. 468–472.

Lancaster, P., and Salkauskas, K., 1986, *Curve and Surface Fitting: an Introduction*, Academic Press, London.

Meinguet, J., 1979, "Multivariate interpolation at arbitrary points made simple," *Z. Angew. Math. Phys.*, Vol. 30, pp. 292–304.

Mortenson, M. E., 1985, *Geometric Modeling*, Wiley, Toronto.

Narcowich, F. J., and Ward, J. D., 1992, "Norm estimates for the inverses of a general class of scattered-data radial-function interpolation matrices," *J. Approx. Theory*, Vol. 69, pp. 84–109.

Ronsky, J. L., 1994, "In-vivo patellofemoral joint contact," Ph.D. Dissertation, University of Calgary Press, Calgary, AB.

Ronsky, J. L., Boyd, S. K., Lichti, D. D., Chapman, M. A., and Salkauskas, K., 1999, "Precise measurement of cat patellofemoral joint geometry with multi-station digital photogrammetry," ASME JOURNAL OF BIOMECHANICAL ENGINEERING, Vol. 121, pp. 196–205.

Scherrer, P. K., and Hillberry, B. M., 1979, "Piece-wise mathematical representation of articular surfaces," *J. Biomechanics*, Vol. 12, pp. 301–311.

Scherrer, P. K., Hillberry, B. M., and Van Sickle, D. C., 1979, "Determining the in-vivo areas of contact in the canine shoulder," ASME JOURNAL OF BIOMECHANICAL ENGINEERING, Vol. 101, pp. 271–278.

Schut, G. H., 1976, "Review of interpolation methods for digital terrain models," *The Canadian Surveyor*, Vol. 30, No. 5, pp. 389–412.

Seedhom, B. B., Longton, E. B., Wright, V., and Dowson, D., 1972, "Dimensions of the knee," *Ann. Rheum. Dis.*, Vol. 31, pp. 54–58.

Shiba, R., Sorbie, C., Siu, D. W., Bryant, J. T., Cooke, D. V., and Wevers, H. W., 1988, "Geometry of the humeroulnar joint," *J. Orthop. Res.*, Vol. 6, pp. 897–906.

Soslowsky, L. J., Flatow, E. L., Bigliani, L. U., Pawluk, R. J., Ateshian, G. A., and Mow, V. C., 1992, "Quantitation of in situ contact areas at the glenohumeral joint: a biomechanical study," *J. Orthop. Res.*, Vol. 10, pp. 524–534.

Wahba, G., 1990, "Spline Models for Observational Data," *Series in Applied Mathematics*, Vol. 59, SIAM.

Wendland, H., 1995, "Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree," *Advances in Computational Mathematics*, Vol. 4, pp. 389–396.

Wismans, J., Veldpaus, F., Janssen, J., Huson, A., and Struben, P., 1980, "A three-dimensional mathematical model of the knee-joint," *J. Biomechanics*, Vol. 13, pp. 667–685.

Wijk, M. C. van, 1980, "Moiré contourgraphy—an accuracy analysis," *J. Biomechanics*, Vol. 13, No. 7, pp. 605–613.