ELSEVIER

Journal of Biomechanics 33 (2000) 901–907

JOURNAL OF BIOMECHANICS

www.elsevier.com/locate/jbiomech

Technical note

The use of basis functions in modelling joint articular surfaces: application to the knee joint

Yasin Y. Dhaher[a,b,c,*,1], Scott L. Delp[d,2], William Z. Rymer[a,b,c]

[a]*Sensory and Motor Performance Program, Rehabilitation Institute of Chicago, 345 East Superior Street, Room 1406, Chicago, IL, 60611, USA*
[b]*Departments of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL, USA*
[c]*Department of Biomedical Engineering, Northwestern University, Chicago, IL, USA*
[d]*Biomechanical Engineering Division, Mechanical Engineering Department, Stanford University, Stanford, CA, USA*

Accepted 18 November 1999

Abstract

This article introduces a new method to represent bone surface geometry for simulations of joint contact. The method uses the inner product of two basis functions to provide a mathematical representation of the joint surfaces. This method guarantees a continuous transition in the direction of the surface normals, an important property for computation of joint contact. Our formulation handles experimental data that are not evenly distributed, a common characteristic of digitized data of musculoskeletal morphologies. The method makes it possible to represent highly curved surfaces, which are encountered in many anatomical structures. The accuracy of this method is demonstrated by modeling the human knee joint. The mean relative percentage error in the representation of the patellar track surface was 0.25% (range 0–1.56%) which corresponded to an absolute error of 0.17 mm (range 0–0.16 mm).

Keywords: Basis functions; Surface modeling; Surface normal; Joint contact

1. Introduction

Accurate mathematical descriptions of articular surfaces are essential in biomechanical models because surface errors can propagate to generate large errors in kinetic and kinematic simulations. To avoid such errors, both geometric and natural compatibility conditions should be met in simulations that include joint contact. The geometric compatibility condition assures that the distance between the contact points on the two surfaces is zero (i.e., penetration is avoided). The natural compatibility condition assures that the normals to the surfaces at the point of contact are collinear. Failure to adequately address these conditions can compromise the accuracy of biomechanical simulations that include surface contact.

Three methods are available that provide mathematical description of articular surfaces. The first method uses two-dimensional polynomials to represent bone surface geometry (Wismans et al., 1980; Engin and Moeinzadeh, 1982; Blankvoort et al., 1991; Tumer and Engin, 1993). While two-dimensional polynomials are reasonably accurate, small errors in the surface fit can create penetration of the two objects in contact and result in multiple contact points. This results in trajectories of the contact point locations that are discontinuous with joint motion. The second method uses polygons to represent bone surfaces (Delp and Loan, 1995). Although polygonal models of bone surfaces provide an acceptable visual representation of bone geometries, the natural compatibility condition is not necessarily satisfied. Since the contact polygons from the two surfaces are not necessarily coplaner, collinearity of the normals at the contact point is not guaranteed. The third method, parametric surface patches, is frequently used to represent bone surfaces (Ateshian et al., 1991; Hirokawa, 1991). As described in Scherrer and Hillberry (1979), each of the

* Corresponding author. Tel.: + 1-312-238-1408; fax: + 1-312-238-2208.

E-mail address: y-dhaher@nwu.edu (Y.Y. Dhaher).

[1] Current address: Sensory and Motor Performance Program, Rehabilitation Institute of Chicago, 345 E. Superior Street, Room 1406, Chicago IL 60611, USA.

[2] Current address: Mechanical Engineering Department, Terman 550, Stanford University, Stanford, CA 94305-3030, USA.

0021-9290/00/$ - see front matter
PII: S0021-9290(00)00024-5

surface patches is bounded by four different cubic splines. Coefficients of the cubic splines for all patches are computed so any two adjacent patches are continuous in the zeroth derivative, the first derivative, and in the cross-derivatives at each of the common corner points. There is no error in the representation at the corners; however, errors are possible for points at other locations on a patch (Scherrer, 1977). For example, the cross-derivatives are usually set to zero (Hart, 1974; Almond, 1991), which degrades the accuracy of representation within a patch. Since there is no guarantee that the errors are uniform within a single patch, or across patches, the continuity of the tangent vectors is not guaranteed. Thus, evaluation of the normal vector (computed as the cross product of the two tangent vectors) at two successive points would not produce a continuous transition in the direction of the surface normals, resulting in a discontinuous contact points in simulations of joint contact (Hefzy and Yang, 1993).

Given the limitations of these previous approaches, Ateshian (1993) introduced the use of B-splines to create geometric models of articular surfaces. The control points of the B-spline were computed from a rectangular grid of points. Any missing points not specified within a rectangular grid were computed with an extrapolation method (Akima, 1978). However, as Akima (1978) acknowledges, the accuracy of this extrapolation grid is questionable. A B-spline formulation that eliminates the use of Akima's method was presented by Ateshian (1995). This important improvement is not without limitations: the method may not represent highly curved surfaces and requires that the digitized anatomy be regularly distributed. In the present paper, we build on this background and attempt to circumvent these limitations by computing the control points of the B-splines from data that are not necessarily placed in a rectangular grid. We also expand the formulation to represent highly curved surfaces, such as the patellar track. Our method insures continuous transition in the direction of the surface normals (i.e., natural compatibility) even when a slight error is associated with the representation.

2. Methods

Our method employs the products of basis functions to represent three-dimensional surfaces, similar to the mathematical formulation proposed in Hayes and Halliday (1974). The basis functions used here, B-splines, were chosen to provide a representation that is continuous to the second derivative and is easy to evaluate. The form of the B-splines chosen is based on the early work of Halliday et al. (1972). Our formulation solves the equations by singular value decomposition, which is well suited to situations in which the experimental data are not conveniently distributed. We address the limitation of the B-splines to model highly curved surfaces by partitioning the data to insure that the functions are single valued and then providing a formulation that guarantees continuity across the partitioning boundary.

The knee joint articular surfaces were derived from the digitized transverse contours of the visible human right knee joint (Visible Productions, Ft. Collins, Co.). A local coordinate system was established for each of the bones. The origins of the femoral and tibial coordinate systems were located at the midpoint of the line between the medial and lateral femoral epicondyles and the medial lateral tibial condyles, respectively. The origin of the patellar coordinate system was located at the centroid of the patella. In all three coordinate systems, the medial (x), posterior (y), and superior (z) directions were chosen to be positive and were assumed to be parallel to a global coordinate system. In the global coordinate system, the (x, z) plane was defined as the coronal plane and the (y, z) plane was defined as the sagittal plane. Digitized data of the articular surfaces were normalized to local anatomical dimensions of the different bones (Mensch and Amstutz, 1975). Although all the contact surfaces of the knee joints (the tibiofemoral and patellofemoral joints) were fitted using our method, the femoral surface of the patellofemoral joint surface will be used to illustrate the surface-fitting problem.

The geometry of the patellar track was partitioned into two sub-surfaces ($\Omega^{(x,y)}$ and $\Omega^{(x,z)}$) using a plane defined as $z = Z_c$ (Fig. 1). Partitioning the surface geometry into two meshes insures a single-valued representation in two

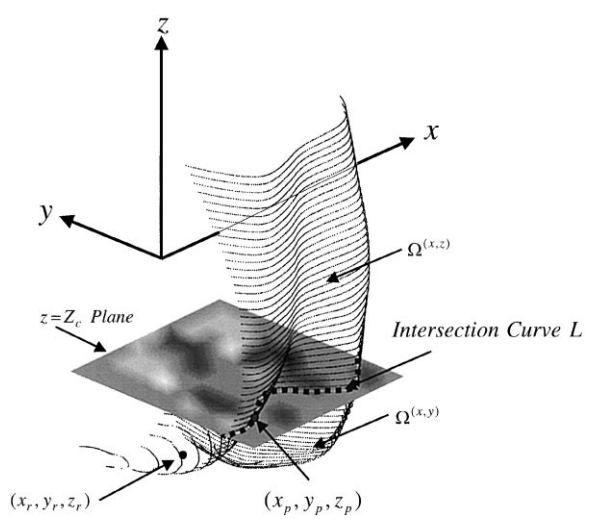

Fig. 1. Representation of the patellar track geometry. The femoral coordinate system (x, y, z) is constructed such that the origin of the coordinate system is located at the midpoint of the line between the medial and lateral epicondyles. The surface geometry is partitioned into two sub-surfaces ($\Omega^{(x,y)}$ and $\Omega^{(x,z)}$) using a plane defined as $z = Z_r$. The (x_r, y_r, z_r) is an arbitrary point located on the $\Omega^{(x,y)}$ sub-surface of the patellar track. the (x_p, y_p, z_p) is an arbitrary point located on the partitioning curve, L, between the $\Omega^{(x,y)}$ and $\Omega^{(x,z)}$ sub-surfaces.

of the three spatial variables. The choice of the basis functions in the three directions, x, y, and z, is determined such that the function and its first and second derivatives are continuous. The sub-surface $\Omega^{(x,y)}$ is represented in terms of an (x, y) mesh and $\Omega^{(x,z)}$ is represented in terms of an (x, z) mesh. Each mesh is defined in terms of a set of knots α^m in both directions as defined below

$$\alpha^m = [\alpha^m_{-3}\ \alpha^m_{-2}\ \alpha^m_{-1}\ \alpha^m_0\ \alpha^m_1 \cdots \alpha^m_{K_m}\ \alpha^m_{K_m+1}\ \alpha^m_{K_m+2}\ \alpha^m_{K_m+3}]. \quad (1)$$

The superscript (m) is a general identifier for the different directions, x, y, and z, and α^m_1 is chosen such that its elements are monotonically increasing ($\alpha^m_{l+1} > \alpha^m_l$). The subscript (K_m) in Eq. (1) is the total number of knots used in the corresponding direction. Each of the knot values is computed from the previous value with an increment $\Delta^m = \alpha^m_{l+1} - \alpha^m_l$; $l = 0, 1, 2, \ldots, K_m$. The values of α^m_0 and $\alpha^m_{K_m}$ are chosen to be the minimum and maximum data points of the corresponding direction. The sets of augmented knots $\alpha^m_{-3}\ \alpha^m_{-2}\ \alpha^m_{-1}$ and $\alpha^m_{K_m+1}\ \alpha^m_{K_m+2}\ \alpha^m_{K_m+3}$ can be chosen arbitrarily.

The general surface representation, of both sub-surfaces, in terms of sets of basis functions in two different directions takes the form

$$\left\{\Omega^{(x,y)} : z = \sum_{i=1}^{K_x+4}\sum_{j=1}^{K_y+4} C^{(x,y)}_{ij} X_i(x) Y_j(y),\ z \leqslant Z_c\right\}, \quad (2)$$

$$\left\{\Omega^{(x,z)} : y = \sum_{i=1}^{K_x+4}\sum_{j=1}^{K_y+4} C^{(x,z)}_{ij} X_i(x) Z_j(z),\ z > Z_c\right\}, \quad (3)$$

where $X_i(x)$, $Y_j(y)$ and $Z_j(z)$ are basis functions, $C^{(x,y)}_{ij}$ and $C^{(x,z)}_{ij}$ are the corresponding set of coefficients, and K_m $(m = x, y, z)$ are the number of pre-specified knots along the x, y, and z directions, respectively. One possible set of basis functions that would insure C^2 continuity, suggested by Hayes and Halliday (1974), is the third-order B-spline. The general B-spline in the x-direction is given as

$$X_i(x) = \sum_{k=0}^{4} \gamma^x_{ik} (x - \alpha^x_{k-i})^3 H(x - \alpha^x_{k-i}), \quad (4)$$

where

$$\gamma^x_{ik} = \left(\prod_{\substack{r=0\\ r\neq k}}^{4} (\alpha^x_{i-k} - \alpha^x_{i-r})\right)^{-1} \left(\frac{6(\alpha^x_{i-2} - \alpha^x_{i-4})(\alpha^x_i - \alpha^x_{i-2})}{(\alpha^x_i - \alpha^x_{i-4})}\right) \quad (5)$$

$$H(x - \alpha^x_{k-i}) = \begin{cases} 1, & x \geqslant \alpha^x_{k-i}, \\ 0, & x < \alpha^x_{k-i}, \end{cases} \quad (6)$$

where γ^x_{ik} is a set of B-spline coefficients computed from a pre-specified set of knots, α^x_i; $i = 1, \ldots, K_x$, which are pre-computed before the fitting procedure is evaluated. The second term on the right-hand side of Eq. (5) is a normalization factor that insures a maximum B-spline value of one (Halliday et al., 1972). By definition, $X_i(x)$ is zero everywhere except in the range $\alpha^x_{i-4} < x < \alpha^x_i$ (see the $X_{i+4}(x)$ spline given in Fig. 2 as an example). Hence, for any x, say $x_r < \alpha^x_i$, there are only four basis functions of non-zero values at x_r as illustrated in Fig. 3. Forms of B-splines in the other directions can be attained by replacing x with y or z in Eqs. (4) through (6).

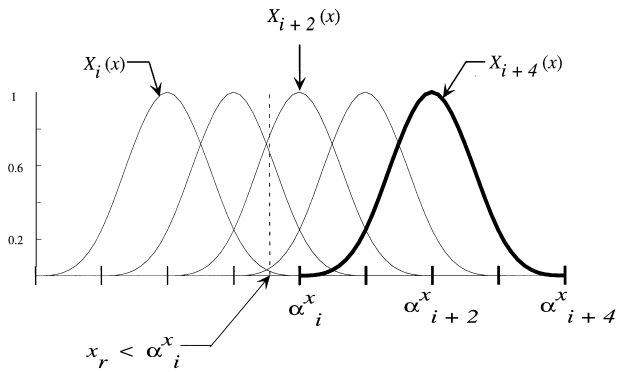

Fig. 2. A typical set of B-splines, $X_i(x)$, $X_{i+2}(x)$, $X_{i+4}(x)$, and an illustrative example for an evaluation of a spline at a given x_r. The set $\alpha^x_i, \alpha^x_{i+1}, \ldots, \alpha^x_{i+4}$ represents the knots that correspond to the $X_{i+4}(x)$ B-spline.

The functional representation of Eqs. (2) and (3) can be used to evaluate the normal vector at any point on the surface. It can be shown that the normal unit vector is given as

$$\boldsymbol{n}^{(x,y)} = \left[\sum_{i=1}^{K_x+4}\sum_{j=1}^{K_y+4} C^{(x,y)}_{ij} \cdot \frac{\mathrm{d}X_i(x)}{\mathrm{d}x} \cdot Y_j(y),\ \sum_{i=1}^{K_x+4}\sum_{j=1}^{K_y+4} C^{(x,y)}_{ij} \cdot X_i(x) \cdot \frac{\mathrm{d}Y_j(y)}{\mathrm{d}y},\ -1\right]^{\mathrm{T}} / |\bullet|, \quad (7)$$

$$\boldsymbol{n}^{(x,z)} = \left[\sum_{i=1}^{K_x+4}\sum_{j=1}^{K_z+4} C^{(x,z)}_{ij} \cdot \frac{\mathrm{d}X_i(x)}{\mathrm{d}x} \cdot Z_j(z),\ -1,\ \times \sum_{i=1}^{K_x+4}\sum_{j=1}^{K_z+4} C^{(x,z)}_{ij} \cdot X_i(x) \cdot \frac{\mathrm{d}Z_j(z)}{\mathrm{d}z}\right]^{\mathrm{T}} / |\bullet|, \quad (8)$$

where $|\bullet|$ represents the magnitude of the numerator vector and the superscript T means 'transpose'. The continuity of the B-spline's first derivatives insures a smooth evaluation of the surface normals given in Eqs. (7) and (8).

Consider, for simplicity, only one of the patellar track sub-surfaces given in Eqs. (2) and (3), say $\Omega^{(x,y)}$. Let z_r be given at points (x_r, y_r), $r = 1, 2, \ldots, n$, which are shown in Fig. 1. The goal is to compute, using Eq. (2), the coefficients $C^{(x,y)}_{ij}$ from the given values of x_r, y_r, and z_r. Substituting each data point into Eq. (2) and using the definition of the B-spline given in Eqs. (4) and (5), the set of the observation equations can be written as follows:

$$\sum_{i=1}^{K_x+4}\sum_{j=1}^{K_y+4} C^{(x,y)}_{ij} X_i(x_r) Y_j(y_r) = z_r; \quad r = 1, 2, \ldots, n. \quad (9)$$

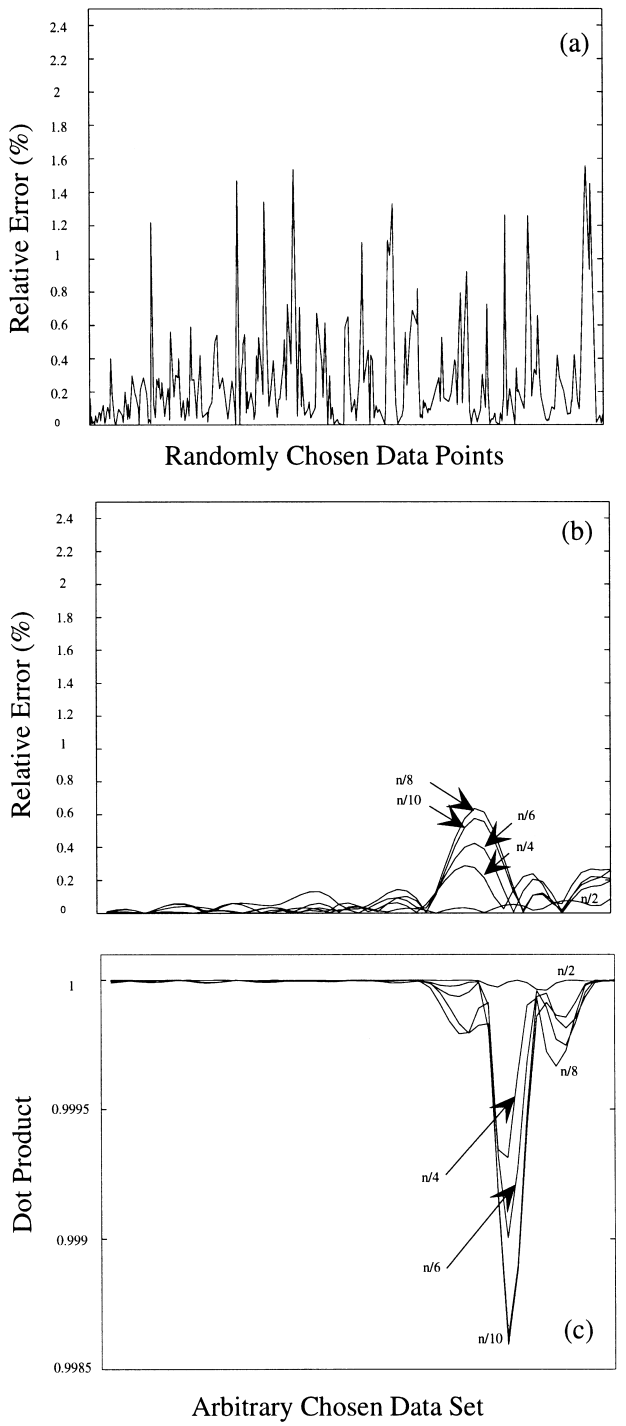

Fig. 3. (a) Relative percentage error versus a randomly selected set of (x, y) points in the $\Omega^{(x,y)}$ sub-space of the patellar track. The relative percentage error was computed with Eq. (14) with $K_x = K_y = 10$, $\Delta^x = 0.0528$, $\Delta^y = 0.0573$, $\alpha_0^x = -0.349$, $\alpha_0^y = 0.621$, $\alpha_{K_x}^x = 0.232$, and $\alpha_{K_x}^y = 0.00899$ used in the calculation of $g^{(x,y)}$ (see text for definitions); (b) Relative percentage error plot versus a randomly distributed set of data points on the patellar track. The relative percentage error was computed with respect to the estimated z values for n data points; (c) The dot product between the normal vectors computed with the full data set (n points) and with half ($n/2$), a quarter ($n/4$), a sixth ($n/6$), an eighth ($n/8$), and a tenth ($n/10$) of the data set evaluated at an arbitrary set of data points. As the relative percentage error increases in (b) the misalignment of the normal also increase.

To compute $C_{ij}^{(x,y)}$ in the above equation, the set of equations are first represented in a matrix form as follows:

$$Ag^{(x,y)} = \bar{z}, \tag{10}$$

where

$$\bar{z} = [z_1, z_2, \ldots, z_n]^{\mathrm{T}}$$

$$g^{(x,y)} = [C_{11}^{(x,y)}, \ldots, C_{1(Kx+4)}^{(x,y)}, C_{21}^{(x,y)}, \ldots, C_{2(Kx+4)}^{(x,y)}, \ldots, C_{(Ky+4)1}^{(x,y)}, \ldots, C_{(Ky+4)(Kx+4)}^{(x,y)}]^{\mathrm{T}}$$

and A contains the values of the inner product $X_i(x)\,Y_j(y)$ for all points in the mesh.

It is clear from the above formulation that A is a rectangular matrix of n rows and $(K_x + 4)(K_y + 4)$ columns. One possible way to solve this problem is to multiply both sides of Eq. (10) by A^{T}, rewriting Eq. (10) as $A^{\mathrm{T}}Ag^{(x,y)} = A^{\mathrm{T}}\bar{z}$. Since, in general, a nonsingular $A^{\mathrm{T}}A$ is not guaranteed, we propose a least-squares solution of Eq. (10) that minimizes $\|Ag^{(x,y)} - \bar{z}\|_2$ and is given by

$$\hat{g}^{(x,y)} = A^{+}\bar{z}, \tag{11}$$

where A^{+} is the generalized inverse of A. The generalized inverse of A can be represented in terms of M, Λ, and N as follows (Ortega, 1987):

$$A^{+} = N\Lambda M^{\mathrm{T}} \tag{12}$$

$$\Lambda = \begin{bmatrix} \mathrm{diag}\,(1/\sigma_1, 1/\sigma_2, \ldots, 1/\sigma_1) & 0 \\ 0 & 0 \end{bmatrix} \in \mathscr{R}^{(K_x+4)(K_y+4)\times n}, \tag{13}$$

where l is the rank of A, $\sigma_1 > \sigma_2 > \cdots > \sigma_l > 0$, are the singular values of the matrix A, and $N \in \mathscr{R}^{(K_x+4)(K_y+4)}$ and $M \in \mathscr{R}^{n}$ are unitary matrices and are computed from the eigenvalues and the associated orthonormal eigenvectors of $A^{\mathrm{T}}A$ (Ortega, 1987). If the rank of $A = (K_x + 4)(K_y + 4)$, Eq. (10) has a unique solution and $A^{+} = A^{-1}$. If the rank of $A < (K_x + 4)(K_y + 4)$, the solution given by Eq. (11) is the best solution of Eq. (10) in a least-squares sense.

To test the accuracy of the model, the data set was randomly divided into two sets. The first set was used to compute the model while the second set was used to compare the model-predicted positions to the actual digitized data. The random selection of data points is also used to simulate the irregular distribution of the experimental data encountered in actual experimental set-ups.

Specifically, after computation of $C_{ij}^{(x,y)}$, the accuracy of the basis function representation was tested by comparing the estimated values with the measured values for all the points in the set that were not used in the calculations of $C_{ij}^{(x,y)}$. To illustrate this, (x, y) mesh of the patellar track

surface (Fig. 1) was used and the relative percentage error in the estimation of the z coordinates of points in the (x, y) mesh was computed with the following equation:

$$\text{Relative Percentage Error} = \frac{|z_r - \hat{z}_r|}{|z_r|} \times 100, \tag{14}$$

where z_r is the actual z-coordinate and $\hat{z}_r$ is the corresponding estimated z-coordinate computed with Eq. (2)

The dependence of the mathematical representation on the number of points used to create the surface was evaluated. The increase in relative percentage error and the change in the direction of the surface normal were computed as the number of points used to create the surface (n) was reduced. The number of points was reduced to $n/2$, $n/4$, $n/6$, $n/8$, and $n/10$. Without loss of generality, the error was evaluated for a set of arbitrary data points, (x_a, y_a, z_a) that lie in the $\Omega^{(x,y)}$ sub-surface of the patellar track geometry (Fig. 1). The procedure starts with the computation of $C_{ij}^{(x,y)}$, following the method described above. For each of the computed $C_{ij}^{(x,y)}$, estimated values of the z-coordinates of the arbitrary data set, $\hat{z}_a$, were obtained using Eq. (2). The associated normal vectors were also computed for each $C_{ij}^{(x,y)}$ using Eq. (7). The estimated positional data, $\hat{z}_a$, were then compared to the true coordinates z_a. As a measure of smoothness, the dot products of the normal vectors obtained using n data points and fewer points ($n/2$, $n/4$, $n/6$, $n/8$, and $n/10$) were computed.

To insure continuity across the partitioning boundary between the $\Omega^{(x,y)}$ and $\Omega^{(x,z)}$ sub-surfaces, the intersection curve L in Fig. 1, a set of equality constraints are imposed on the calculation of $C_{ij}^{(x,z)}$ of Eq. (3). Since Eq. (2) is valid for all $z \leqslant Z_c$, the continuity of the representation of the $\Omega^{(x,y)}$ sub-surface at $z = Z_c$ is guaranteed. Following the procedure used to compute $C_{ij}^{(x,y)}$, the matrix equation used to calculate $C_{ij}^{(x,z)}$ can be written in the following form:

$$\bar{A} g^{(x,z)} = \bar{y}, \tag{15}$$

where

$$\bar{y} = [y_1, y_2, \ldots, y_{n_l}]^{\mathrm{T}}$$

$$g^{(x,z)} = [C_{1l}^{(x,z)}, \ldots, C_{1(Kx+4)}^{(x,z)}, C_{2l}^{(x,z)}, \ldots, C_{2(Kx+4)}^{(x,z)}, \ldots, C_{(Kz+4)l}^{(x,z)}, \ldots, C_{(Kz+4)(Kx+4)}^{(x,z)}]^{\mathrm{T}}$$

and $\bar{A}$ contains the values of the inner product $X_i(x) \cdot Z_j(z)$ for all points in the mesh, (x_l, y_l, z_l), $l = 1, 2, \ldots, n_l$. If $(x_p, y_p, z_p = Z_c)$, $p = 1, 2, \ldots, k$, is the set of data points at the intersection curve L between the two sub-surfaces, then the coefficient vector $g^{(x,z)}$ can be computed subject to the following constraint:

$$\bar{\bar{\mathbf{A}}} g^{(x,z)} = \bar{\bar{\mathbf{y}}}, \tag{16}$$

where

$$\bar{\bar{\mathbf{y}}} = [y_1, y_2, \ldots, y_k]^{\mathrm{T}}$$

$$g^{(x,z)} = [C_{1l}^{(x,z)}, \ldots, C_{1(Kx+4)}^{(x,z)}, C_{2l}^{(x,z)}, \ldots, C_{2(Kx+4)}^{(x,z)}, \ldots, C_{(Kz+4)l}^{(x,z)}, \ldots, C_{(Kz+4)(Kx+4)}^{(x,z)}]^{\mathrm{T}}$$

and $\bar{\bar{\mathbf{A}}}$ contains the values of the inner product $X_i(x) \cdot Z_j(z)$ for all points on the intersection curve $L, (x_p, y_p, z_p = Z_c)$, $p = 1, 2, \ldots, k$.

By the use of Lagrange multipliers, λ, the required solution of Eq. (15) subject to Eq. (16) requires the solution to the equation

$$\underbrace{\begin{bmatrix} \bar{A} & \vdots & \bar{\bar{\mathbf{A}}} \\ \cdots & \vdots & \cdots \\ \bar{\bar{\mathbf{A}}} & \vdots & 0 \end{bmatrix}}_{B} \begin{bmatrix} g^{(x,z)} \\ \cdots \\ \lambda \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \cdots \\ \bar{\bar{\mathbf{y}}} \end{bmatrix}. \tag{17}$$

A least-squares solution of the above equation can be obtained through the computation of the generalized inverse of B (see Eqs. (11)–(13).

3. Results

Basis functions provided a highly accurate representation of the patellar track geometry. The relative percentage error between the estimated and the actual z values was always less than 1.56% (Fig. 3a). The mean relative percentage error was 0.25% over all the data points used in the fitting. The increase in the error was less than 0.09% when only half of the data points were used in the fit (Fig. 3b). The relative percentage error increased as fewer points were used to create the surface. However, the relative percentage error increased by only 0.9% when as few as one tenth of the data points were used. The increase in the error was greater in regions with high curvature. In general, the misalignment of the surface normals increased as fewer points were used (Fig. 3c). However, even when as few as one tenth of the data points were used, the angle between the normal vectors was less than 3.2°. This shows that the method provides an accurate representation even when fewer data points are used to construct the mathematical function of the surface.

Similar results, not shown, were obtained for the rest of the contact surfaces of the knee. The mean percentage error was 0.15% (range 0–0.53%) over all the data points used in the fitting of the medial and lateral tibial plateau and a mean percentage error of 0.16% (range 0–0.95%) over all the data points used in the fitting of the patellar surface.

The method provided an accurate representation at the partitioning boundary of the highly curved surface of

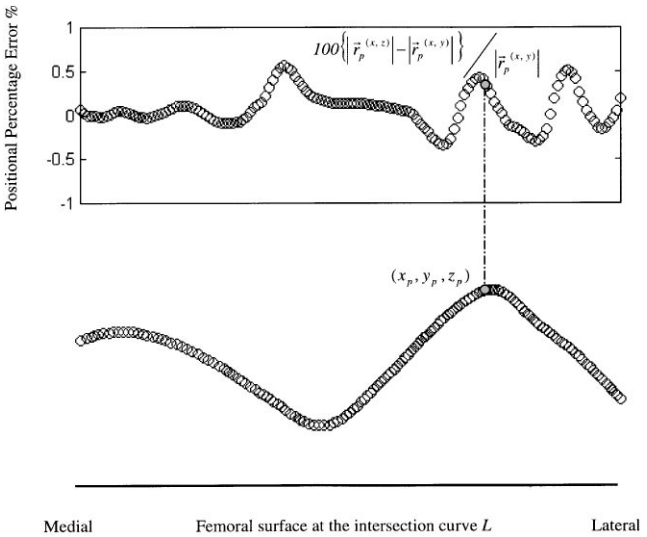

Fig. 4. A horizontal view of the partition curve L, between the two sub-surfaces ($\Omega^{(x,y)}$ and $\Omega^{(x,z)}$) and the relative percentage error of the position vector of all the points (x_p, y_p, z_p) along L measured with respect to the origin of the femoral coordinate system. All position vectors were referenced to the origin of the femoral coordinate system. $\boldsymbol{r}_p^{(x,y)}$ and $\boldsymbol{r}_p^{(x,z)}$ are the position vectors of the point (x_p, y_p, z_p) computed from the fitting of the $\Omega^{(x,y)}$ and $\Omega^{(x,z)}$ sub-surfaces, respectively. Parameters used to estimate the $g^{(x,z)}$ vector using Eq. (17) were identical to the parameters used in the computation of $g^{(x,z)}$ with $\Delta^x = \Delta^y$ and $K_x = K_y$ (see the legend of Fig. 3 for numerical values).

the patellar track (Fig. 4). The mean relative percentage error was 0.07% (range 0.34%–0.56%) over all the data points along the partitioning curve L.

4. Discussion

In this study, we used basis functions to represent joint surface geometry. This method accurately described the knee joint contact surfaces when compared to the original data set. Furthermore, the singular-value decomposition method used here insures accurate fitting in cases where the data points are irregularly distributed. Accurate surface models were obtained with a small subset of the original data points. The current method also insures continuity of the normal vector field, even when errors are present. In addition, we have shown that accurate representation of highly curved surfaces is possible.

While offering some advantages, there are several requirements of the method. First, the method assumes that the surface is a single-valued function of two of the three spatial variables. In the case of a highly curved surface, and in order meet this condition, the data points may need to be partitioned as they were in the example presented here. Once partitioning of the surface is established, additional constraints need to be added to insure continuity along the partitioning curve as they were in the example considered here. Second, there are constraints on the number of data points to be included. Specifically, to avoid the rank-deficiency of the A matrix in Eq. (10), the number of data points (n) should at least be equal to $(K_x + 4)(K_y + 4)$ where K_x and K_y represent the number of knots along the directions of the independent variables x and y, respectively. If n were less than $(K_x + 4)(K_y + 4)$ the system of equations given in Eq. (10) is underdetermined, since there are more unknowns than equations; in this case, the system of equations would have nontrivial solutions only if the system were consistent (Ortega, 1987).

Models used to simulate the dynamics of joints routinely include articular contact constraints. When objects in contact are assumed to be rigid (Wismans, 1980), a point contact assumption is commonly used. If the contact surfaces include some irregularities, small movements of the contact point may result in an abrupt change in the direction of the contact force. Because the contact force lies in the direction of the normal to the surface vector (in a friction-free contact), rapid changes in the surface normal can degrade the estimates of kinetic equilibrium (Hirokawa, 1992). This problem is avoided when surfaces are represented with basis functions because the normal vector field is continuous.

Acknowledgements

The authors gratefully acknowledge the Dr. Ralph and Marian C. Falk Medical Research Trust and the National Institute of Health (RO1 HD33929) for their support of this research.

References

Almond, D.B., 1991. Numerical control for machining complex surfaces. IBM Systems Journal, 150–168.

Ateshian, G.A., 1993. A B-spline least-squares surface-fitting method for articular surfaces of diarthrodial joints. Journal of Biomechanical Engineering 115, 366–373.

Ateshian, G.A. 1995. Generating trimmed B-spline models for articular cartilage layers from uniformed 3D surface data points. In the Proceeding of the Bioengineering Conference, American Society of Mechanical Engineers, BED-Vol. 29, pp. 217–218.

Ateshian, G.A., Soslowsky, L.J., Mow, V.C., 1991. Quantitation of articular surface topography and cartilage thickness in knee joints using strereophotogrammetry. Journal of Biomechanics 24, 761–776.

Blankvoort, L., Kuiper, J.H., Huiskes, R., Grootenboer, H.J., 1991. Articular contact in a three-dimensional model of the knee. Journal of Biomechanics 24, 1019–1031.

Delp, S.L., Loan, J.P., 1995. A graphic-based software system to develop and analyze models of musculoskeletal structures. Computers in Biology and Medicine 25, 21–34.

Engin, A.E., Moeinzadeh, M.H., 1982. Modeling of human joint structure, AFAMRL Report, AFAMRL-TR-81-117.

Halliday, J, Wall, J.F., Joyner, W.D., 1972. Report on multivariable curve fitting using fundamental splines. Maths Services Note 167, British Aircraft Corp., Preston.

Hart, W.B., 1974. Computer graphics aid fuselage design. The least-squares. Engineering Materials and Design Institute, 205–209.

Hayes, J.G., Halliday, J., 1974. The least-squares fitting of cubic spline surfaces to general data sets. Journal of Institute of Mathematics and Applications 14, 89–103.

Hefzy, M.S., Yang, H., 1993. A three-dimensional anatomical model of the human patellofemoral joint, for the determination of the patello-femoral motion and contact characteristics. Journal of Biomechanical Engineering 15, 289–302.

Hirokawa, S., 1991. Three-dimensional mathematical model analysis of the patellofemoral joint. Journal of Biomechanics 24, 659–671.

Hirokawa, S., 1992. Effects of variation on extensor elements and operative procedures in patellofemoral disorders. Journal of Biomechanics 25, 1393–1401.

Mensch, J.S., Amstutz, H.C., 1975. Knee morphology as a guide to knee replacement. Clinical Orthopaedics 112, 231–241.

Ortega, M., 1987. Matrix Theory — A Second Course. Plenum Press, New York.

Scherrer, P.K., 1977. Determining the distance between the surface of two rigid bodies using parametric surface patches and an instrumental spatial linkage. Ph.D. Thesis. Purdue University.

Scherrer, P.K., Hillberry, B.M., 1979. Piecewise mathematical representation of articular surfaces. Journal of Biomechanics 12, 301–311.

Tumer, S.T., Engin, A.E., 1993. Three-Body segment dynamic model of the human knee. Journal of Biomechanical Engineering ASME Transactions 115, 350–356.

Wismans, J., Veldpaus, F., Janssen, J., Huson, A., Struben, P., 1980. A three-dimensional mathematical model of the knee-joint. Journal of Biomechanics 13, 677–685.