COMPUTER GRAPHICS AND IMAGE PROCESSING 17, 145-160 (1981)

Local Invariants and Twist Vectors in Computer-Aided Geometric Design

S. A. SELESNICK

Department of Mathematical Sciences, University of Missouri, St. Louis, Missouri 63121

Received June 3, 1980; revised February 10, 1981

A new method is suggested for choosing the "twist" parameters (or mixed partial derivatives), which arise when Coons patches are used for surface definition or interpolation over rectangular arrays of data points. The method is based on the observation that the component of the twist normal to the surface may be related to a geometrical invariant intrinsic to the surface, namely, the Gaussian curvature. This quantity can also be expressed independently of the twist and estimated on the basis of independent univariate data already available once the patch boundaries are specified (thus avoiding the incompatibilities inherent in bivariate discretizations). The tangential components of the twist vector may also be estimated by independent univariate means from prior data. The resulting surface has the given or estimated curvature at the data points. (Smoothness properties of certain composite surfaces are also analyzed in terms of invariants related to curvature.)

0. INTRODUCTION

The problem of specifying the mixed partial derivative vectors that inevitably arise in parametric surface interpolation schemes over rectangles is an old and vexing one (see [1; 2, Chap. 7] and their references). Despite the terminology, the naive geometrical view that these vectors, being derivatives of tangent vectors, should represent a "twisting" of these tangents and hence give rise to explicable surface effects, does not lend itself to a consistent interpretation. Thus it is possible to parameterize a plane in such a way that these "twists" are everywhere nonzero [2, p. 203] and, conversely, it is possible to locally parameterize any sufficiently smooth surface (at least in a neighborhood of an ordinary point) in such a way that the twist vanishes (see the final remark in Section 3), so that little intrinsic geometrical significance can be ascribed to the vanishing of this vector. On the other hand, approximations to the derivatives by means of finite differences of discretely sampled data cannot guarantee the compatibilities required by even mild continuity assumptions and consequently such direct methods require elaborate corrective action to be taken. These have been successfully developed. Less direct methods involving the interpolation of intermediate surfaces have also been applied (see [1] for a discussion of these). Another successful method, involving assumptions of $C^{2,2}$ continuity, is described in detail in [2, Sect. 7.2.1].

In this note we propose an alternative approach to the problem which is based upon the consideration of certain intrinsic geometrical invariants. We have noted that the twist vector in isolation cannot directly reflect intrinsic geometrical properties of the surface involved; nevertheless, there is a significant invariant, namely, the Gaussian curvature, in which the square of the normal component of this vector participates. The classical "theorema egregium" of Gauss (see [3, 4]) asserts not only that the curvature is independent of any parameterization but also that it is independent of the normal embedding of the surface into Euclidean 3-space: thus it should be possible to calculate this quantity from data intrinsic to the surface (i.e.,

0146-664X/81/100145-16$02.00/0

the "metric"). Knowledge of the Gaussian curvature thus serves (in most cases) to determine the normal component of the twist up to a (resolvable) ambiguity in sign. The components of the twist along the tangent vectors (at a point) submit more readily to a calculable interpretation and these three quantities specify the twist vector completely, at least at ordinary points. This method avoids the usual incompatibilities arising from the sampling of discrete bivariate data. (A consideration of Gaussian curvature also sheds some light on two curious effects which have been observed in surfaces produced by rectangular patch-wise interpolation. These concern the effects produced, respectively, by assigning zero twists at patch corners (see Section 3), and by using nonuniform parameter intervals for knot spacing along the boundary splines of composite surfaces (see Section 2).)

The contents of the paper are as follows. In Section 1 we describe an algorithm to calculate the twist vector at a point once the Gaussian curvature and certain univariate quantities are known. An algorithm for estimating the Gaussian curvature in terms of readily available first order information is described. In Section 2 we consider some properties of composite surfaces which are entailed by a consideration of an invariant closely related to the Gaussian curvature, namely, the Weingarten map. Finally (Section 3) there is a mathematical Appendix in which most of the results used are derived from first principles.

We consider only bicubic patches—appropriate generalizations to more general schemes will be apparent.

1. THE BICUBIC PATCH

A "bicubic patch" is the image of a function

$$\mathbf{r}\colon [a, b]^2 \to \mathbb{R}^3$$

of the form

$$\mathbf{r}(u, v) = \sum_{i,j=0}^{3} \mathbf{a}_{ij} u^i v^j$$

with $(u, v) \in [a, b]^2$, and where the $\mathbf{a}_{ij}$ are constant 3-vectors. Conventionally one takes $a = 0$, $b = 1$ in which case the 16 constant vectors can be interpreted in terms of corner data. Specifically, the last equation can be rearranged and expressed in the form

$$\mathbf{r}(u, v) = F(u)\, Q\, F(v)^T, \tag{1.1}$$

where $F(u)$ is a 4-vector whose entries are the Hermite cubic basis polynomials

$$F(u) = (1 - 3u^2 + 2u^3, 3u^2 - 2u^3, u - 2u^2 + u^3, -u^2 + u^3)$$

and Q is the 4×4 matrix of 3-vectors

$$\begin{pmatrix} \mathbf{r}(0,0) & \mathbf{r}(0,1) & \mathbf{r}_2(0,0) & \mathbf{r}_2(0,1) \\ \mathbf{r}(1,0) & \mathbf{r}(1,1) & \mathbf{r}_2(1,0) & \mathbf{r}_2(1,1) \\ \mathbf{r}_1(0,0) & \mathbf{r}_1(0,1) & \mathbf{r}_{12}(0,0) & \mathbf{r}_{12}(0,1) \\ \mathbf{r}_1(1,0) & \mathbf{r}_1(1,1) & \mathbf{r}_{12}(1,0) & \mathbf{r}_{12}(1,1) \end{pmatrix}.$$

(This is a special case of a Coons tensor-product patch; see [2, Sect. 7.2].) As we have mentioned above, the problem has traditionally resided in the assignment of the twist vectors occupying the lower right-hand 2×2 submatrix of Q, since these quantities, unlike the other data required by Q, do not admit a reliable intuitive geometric interpretation. Moreover, being bivariate quantities, their direct approximation, via finite differences say, is bound to lead to incompatibilities. We shall show that once the tangents in Q have been assigned, notions of Gaussian curvature may be used to reduce the problem of estimating the twist vectors to that of estimating certain other independently univariate quantities. First we introduce some standard notation. For an arbitrary (sufficiently smooth) surface described by a radius vector $\mathbf{r}(u, v)$, the functions E, F, G of (u, v) are defined to be, respectively, $\mathbf{r}_1^2, \mathbf{r}_1 \cdot \mathbf{r}_2, \mathbf{r}_2^2$ where the subscript 1 (resp. 2) denotes partial differentiation with respect to u (resp. v) and $\mathbf{r}_1^2 = \mathbf{r}_1 \cdot \mathbf{r}_1$ (etc.). We also write

$$H^2 = EG - F^2,$$

which is positive at points on the surface for which the normal vector

$$\mathbf{n} = \mathbf{r}_1 \wedge \mathbf{r}_2$$

(the wedge denoting cross product) is well defined—the so-called *ordinary* points. We observe that

$$H^2 = \mathbf{n}^2.$$

The functions L, M, N are defined to be, respectively, $\hat{n} \cdot \mathbf{r}_{11}$, $\hat{n} \cdot \mathbf{r}_{12}$, $\hat{n} \cdot \mathbf{r}_{22}$, where $\hat{n}$ denotes the *unit* normal vector, that is,

$$\hat{n} = \mathbf{n}/|\mathbf{n}|.$$

In terms of these quantities, the Gaussian curvature K (at a point on the surface corresponding to parameter values (u, v)) may be expressed by the following formulae, the first of which is standard (see Section 3 for a derivation of both of these formulae).

$$\begin{aligned} K &= LN - M^2/H^2 \\ &= \frac{(\hat{n} \cdot \mathbf{r}_{11})(\hat{n} \cdot \mathbf{r}_{22}) - (\hat{n} \cdot \mathbf{r}_{12})^2}{\mathbf{r}_1^2 \mathbf{r}_2^2 - (\mathbf{r}_1 \cdot \mathbf{r}_2)^2}, \end{aligned} \tag{1.2}$$

$$K = \frac{\hat{n}_1 \wedge \hat{n}_2 \cdot \mathbf{n}}{\mathbf{n}^2}. \tag{1.3}$$

It is important to note that formula (1.2) explicitly involves the normal component (M) of the twist vector whereas (1.3) does not. We mention also the following remarkable facts concerning this quantity. First, it is independent of the particular (u, v)-parameterization used to describe the surface (this is demonstrated in Section 3). Second, it is independent of the "normal embedding" of the surface into the ambient 3-dimensional Euclidean space. That is to say, despite the appearance of the above formulae, the value of K is in fact *independent* of the information coming from the surface normal. Indeed, K depends only upon the three quantities E, F, G (which determine the so-called "metric"). This is the content of Gauss' "theorema egregium" (see [3, p. 29; 4, p. 114]). Thus the curvature can in principle be supplied at a point if enough is known about the mutual disposition of neighboring points. (A 2-dimensional entity entirely confined to the surface would be able to determine K by making measurements that are entirely confined to the surface.) In the case of practical interest, namely, when only discrete aggregates of point data are available, there is usually enough implied information about the local behavior of the surface normal to enable an estimate of the Gaussian curvature to be made so that the full force of Gauss' theorem is not actually required. Nevertheless, it is important to realize that the Gaussian curvature has intrinsic (indeed topological) significance.

Returning to the bicubic patch (1.1) let us suppose that the tangents have been assigned in the Q matrix. Then at each corner the surface normal, $\mathbf{r}_{11}$, $\mathbf{r}_{22}$, and the quantity H^2 are fully determined. Thus, with the exception of $M(=\hat{n}\cdot\mathbf{r}_{12})$, everything on the right-hand side of (1.2) is determined. If the Gaussian curvature K at a corner were known, (1.2) could be solved to give

$$\hat{n}\cdot\mathbf{r}_{12} = \pm\left((\hat{n}\cdot\mathbf{r}_{11})(\hat{n}\cdot\mathbf{r}_{22}) - KH^2\right)^{1/2}. \tag{1.4}$$

Postponing for a moment the choice of sign, the possibility arises that the expression under the radical may be negative. In this eventuality one would be justified in asserting that there was an obstruction to setting up a bicubic patch having the given curvature and boundary conditions at the corner in question. The geometrical nature of such an obstruction is most clearly seen in the extreme case in which $\hat{n}\cdot\mathbf{r}_{11}$ and $\hat{n}\cdot\mathbf{r}_{22}$ have opposite sign and K is asked to be positive. In other cases one notes that the tangents at the corner in question can be manipulated in the tangent plane without affecting the values of $\hat{n}\cdot\mathbf{r}_{11}$ and $\hat{n}\cdot\mathbf{r}_{22}$. (To see this, merely express the boundary curves in Hermite form and differentiate twice.) Thus the value of H^2 can be altered to give a real value to $\hat{n}\cdot\mathbf{r}_{12}$. In other words, the adjacent patch boundaries can be manipulated to accommodate the given curvature at the corner (in the absence of the sign conflict mentioned above). However, in most cases of practical interest, for instance, when the boundaries are spline segments, one is not at liberty to alter the tangents, and if $\hat{n}\cdot\mathbf{r}_{12}$ threatens to become imaginary in such cases one has no choice but to compromise the value of K. A reasonable choice for the revised value of K would be the one nearest the given value, which yields a real value for $\hat{n}\cdot\mathbf{r}_{12}$. Since the expression under the radical is linear in K, this value is clearly the one making $\hat{n}\cdot\mathbf{r}_{12}$ vanish; i.e., LN/H^2. This simpler option is the one we shall use in this paper.

To help decide the choice of signs in (1.4) we shall argue heuristically on the assumption that the variation of the interpolated surface between the data points should be less than the variation implied by the data points themselves. In this case we would seek to diminish the undulations of the transverse tangents (relative to some mean plane) along each boundary curve, since such undulations would propagate through the patch with little attenuation, owing to the low order of the polynomials defining the patch surface. Now the normal component of the twist is a measure of the rate of this undulation relative to the tangent plane at the point in question, and so to diminish its variation (i.e., the number of turning points) between the data points we should choose this component to have *opposite* signs at the ends of a patch boundary. Thus if we choose the positive sign in (1.4) at the data point corresponding to the parameter values $u = 0$, $v = 0$, the signs required at the other patch corners are determined. In this case it is easy to see that our sign condition is equivalent to the requirement that the interior vertices of the Bézier polyhedron corresponding to the patch should lie *above* the corresponding tangent planes (at the corners) defined by the outer Bézier vertices. (Here "above" means "in the direction of the surface normal.") This kind of convex arrangement is of course the usual configuration chosen for the Bézier polyhedron. (See [2, Sect. 7.2.2] for the correspondence between the Hermite and Bézier forms.)

Having arrived at a value for the normal component of the twist vector at a corner, and assuming that the data point in question is an ordinary point, it suffices in order to specify the twist vector completely, to supply values for the tangential components of this vector. Now these components can be readily interpreted in terms of univariate data. Thus, at a corner point we may write

$$\begin{aligned}\mathbf{r}_1 &= \frac{\partial \mathbf{r}}{\partial u}\\ &= \frac{d\mathbf{r}}{ds}\frac{ds}{du},\end{aligned}$$

where s denotes arc length measured along the boundary $v = \text{constant}$. Thus

$$\begin{aligned}\mathbf{r}_1^2 &= \left[\frac{d\mathbf{r}}{ds}\right]^2\left[\frac{ds}{du}\right]^2\\ &= \left[\frac{ds}{du}\right]^2\end{aligned}$$

since $d\mathbf{r}/ds$ is the unit tangent vector. So, differentiating with respect to v, we have

$$\mathbf{r}_1 \cdot \mathbf{r}_{12} = \left[\frac{ds}{du}\right]\frac{d}{dv}\left[\frac{ds}{du}\right] \tag{1.5}$$

and similarly

$$\mathbf{r}_2 \cdot \mathbf{r}_{21} = \left[\frac{ds}{dv}\right]\frac{d}{du}\left[\frac{ds}{dv}\right] \tag{1.6}$$

(where we are of course assuming that $\mathbf{r}_{12} = \mathbf{r}_{21}$). But from the equation preceding (1.5) we note that ds/du (> 0) is the length of the (assumed known) tangent $\mathbf{r}_1$, and the v-derivative of this length can be estimated by means of some univariate interpolation procedure involving only the variation with v of the lengths ds/du of the known tangents in the u-direction (and similarly for the right-hand side of (1.6)).

Alternatively one could assume that the map

$$\mathbf{r}\colon [0, 1]^2 \to \mathbb{R}^3$$

factors through pieces of surface with given metrics so as to induce local isometries at the corners. That is to say, locally at a corner, the patch can be developed onto some known or preassigned surface in such a way that the (u, v) parameters are isometrically related (e.g., equal) to known parameters for the latter. In other words, we have *local isometric models* for the corners. Such models will then not only determine $d/du(ds/dv)$, etc., but K as well (since by Gauss' theorem this quantity depends only upon the metric), and consequently the entire twist vector.

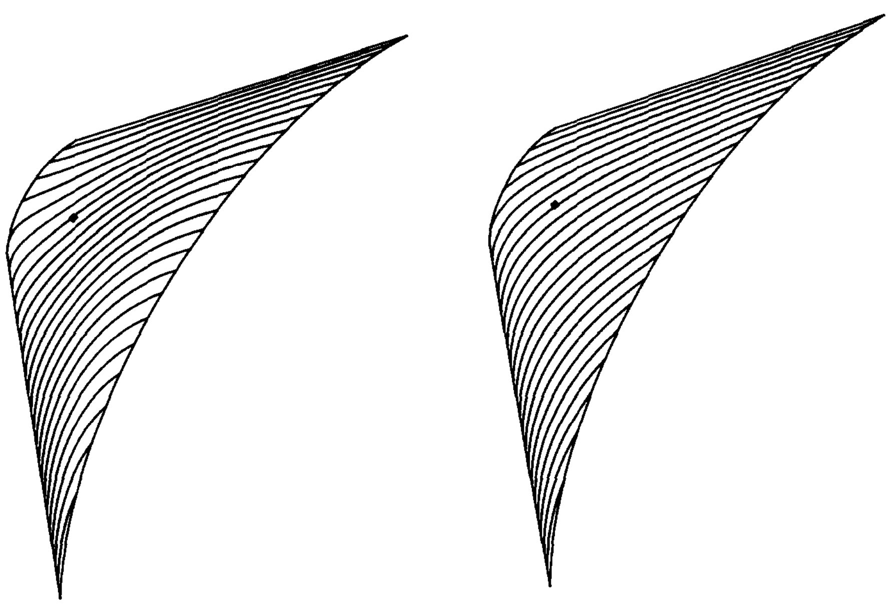
The three equations (1.4)–(1.6) can then be solved for the components of the unknown vector $\mathbf{r}_{12}$, which, when incorporated into (1.1), yields a patch with the given Gaussian curvature at the corner data points (unless the aforementioned pathology arises in (1.4)).

In the following examples, all the curves shown, excepting the boundary curves, are planar sections through the surface. These shapes have been chosen to accentuate the generally subtle effects of changes in the twists; in practice, these surfaces would be modelled with more patches. (In each case the curvature is known and the effect of its imposition is compared with the case of zero twists—for a similar comparison of the effects induced by other schemes, which do not necessarily reproduce the correct curvature at the data points, the reader is referred to [1]).

Examples

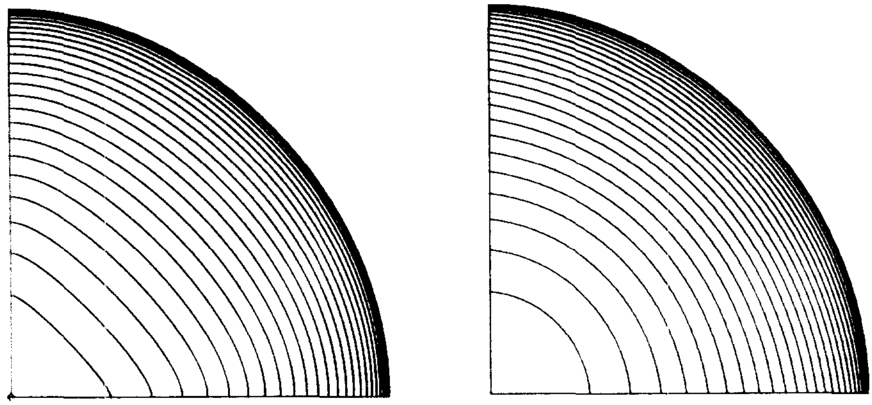
Figure 1. A single bicubic patch in the shape of an approximate conic segment. Figure 1(a) shows the case of zero twists at all corners. Figure 1(b) shows the case of twists calculated as in the text ($K = 0$), with linear interpolation used to estimate the (second) derivatives on the right-hand sides of (1.5), (1.6).

Figure 2. A single degenerate (i.e., three-sided) bicubic patch in the shape of an approximate spherical octant, with the degenerate vertex at the origin of the figure.



a b

FIG. 1. (a) Approximate conic segment over one patch. Zero twists at all corners. Note appearance of curvature inside patch. (b) Approximate conic segment over one patch. Twists calculated as in the text.



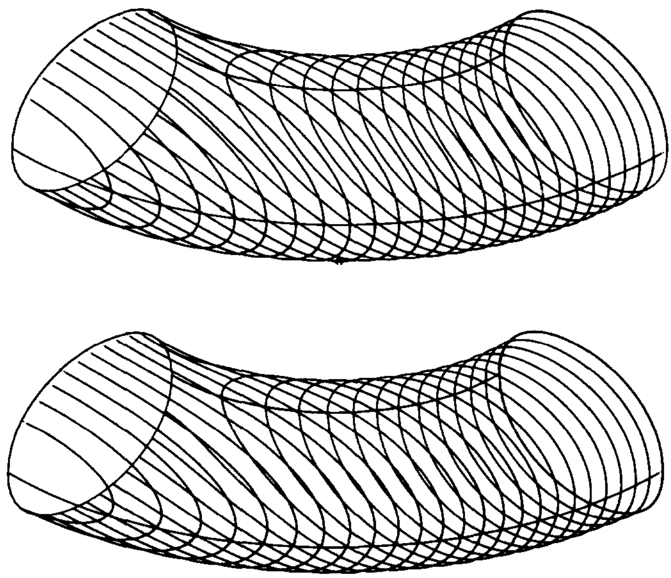
a b

FIG. 2. (a) Approximate spherical octant over one patch, degenerate at the origin. Zero twists at all corners. Note distortion near degeneracy. (b) Approximate spherical octant over one patch, degenerate at the origin. Twists calculated as in the text. Note improvement near degeneracy.

Figure 2(a) shows the case of zero twists at all corners. Note the distortion in the vicinity of the degenerate vertex: this commonly observed phenomenon is a result of the singular nature of the vertex which in a sense concentrates the curvature deviations (induced by the zero twists) which have accumulated at the degenerating vertices as the limit is reached. Figure 2(b) shows the case of twists calculated as in the text, using the idea of local isometric models, the model in this case being based upon spherical polar coordinates. Here the twist at the degenerate vertex has been obtained by going to the limit *through the local isometric model.* In this way the singularity at the degenerate vertex has been resolved, leaving the local spherical topology (largely) intact.

Figure 3. Two bicubic patches in the shape of an approximate toral quadrant. Figure 3(a) shows the case of zero twists at all corners. Note the sagging visible along the envelope of sections through the lower surface. Figure 3(b) shows the case of twists chosen as in the text, using local isometric models. The correct curvature at the patch corners on the inner equatorial boundary yields imaginary values for the twist components. Accordingly these components were taken to be zero, giving zero twists along the inner boundary. (The resulting curvature in this case is about 1.409... times the correct value.) The correct curvature at the patch corners on the outer equatorial boundary yields real nonzero values for the normal twist components. The effect of the resulting twists is to tighten up the surface so that the envelope of sections through the lower surface can be seen to follow the outer equatorial boundary more closely than in the zero twist case. (The quantity $\hat{n} \cdot \mathbf{r}_{11}$ is continuous at the patch corners on the circular boundaries of this composite surface, though $\mathbf{r}_{11}$ itself is not. Consequently our method yields unambiguous twists at these points the surface is C^1-smooth; see Section 2.)

To use the method for general surface fitting, where local models are not available, a method of evaluation or estimating the Gaussian curvature at each data point is clearly required. A simple algorithm is suggested by formula (1.3), concerning which we can make the following observations. The right-hand side of the formula involves only quantities which are available as soon as the patch boundary



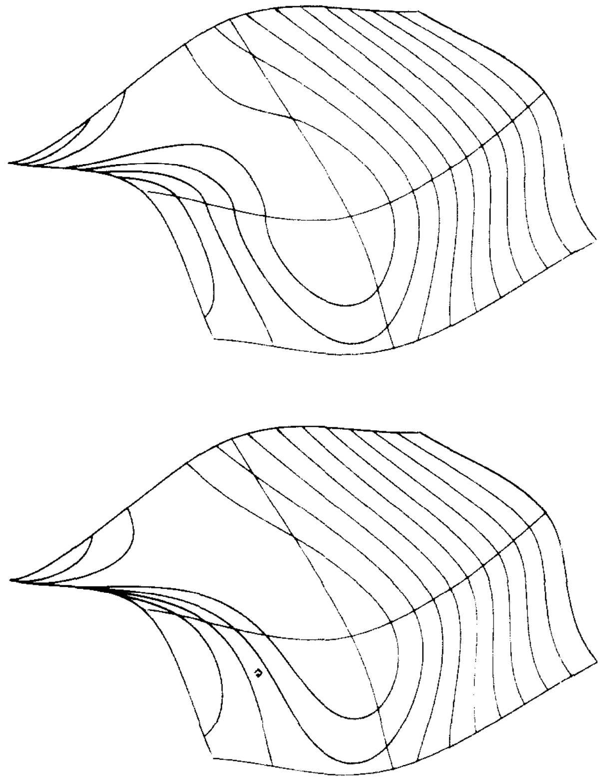
a

b

FIG. 3. (a) Approximate toral quadrant over two patches. Zero twists at all patch corners. Note sagging of lower surface. (b) Approximate toral quadrant over two patches. Twists calculated as in the text. Note tightening of lower surface.

curves have been fixed, and univariate derivatives thereof. That is, once the tangents are given at each data point then **n** and $\hat{n}$ are determined, after which $\hat{n}_1$ and $\hat{n}_2$ can be independently estimated by univariate schemes. For example, one may take at each point running averages of finite differences of the unit normal in the u direction to estimate $\hat{n}_1$ and in the v direction to estimate $\hat{n}_2$. If this is done, the effect is as if an interaction has been introduced between parts of the surface which smooths out the movement of the unit normal (in the manner of a low-pass filter). For a physical surface this effect is consistent with the imposition of a "tension" upon the surface which should be greater than that inherent in the surface if zero twists were supplied, since in this case there would in general be no coupling imposed between the normals. Other physical parameters being equal (and a surface pressure being assumed to act down the normal), this would imply a lowering of the overall mean curvature of the surface (cf. [5, p. 222, Eq. (110)]). Figure 4 illustrates this for a four patch C^1 surface interpolated through arbitrary data. The boundary curves are splines with unit parameter intervals and clamped end conditions. Figure 4(a) shows the case of zero twists and Fig. 4(b) shows the case in which the twists have been chosen according to the method of the text, with curvature estimated by the algorithm just described. (The tangential twist components have also been estimated by running averages of finite differences.) One notes that the latter surface does indeed appear to cling more tightly to the "wire frame," while maintaining C^1 smoothness.

More elaborate curvature algorithms are possible that exploit Gauss' theorem and its generalizations more fully. In addition, the idea of using invariants to build physical properties into the modelled surfaces can also be developed. These topics will be taken up elsewhere.



b

a

FIG. 4. (a) A four patch surface interpolating arbitrary data. Patch boundaries are splines with unit parametrization between knots. Zero twists at all patch corners. (b) A four patch surface as in (a), but with twists chosen according to curvature algorithm described in text.

2. PATCH ASSEMBLIES

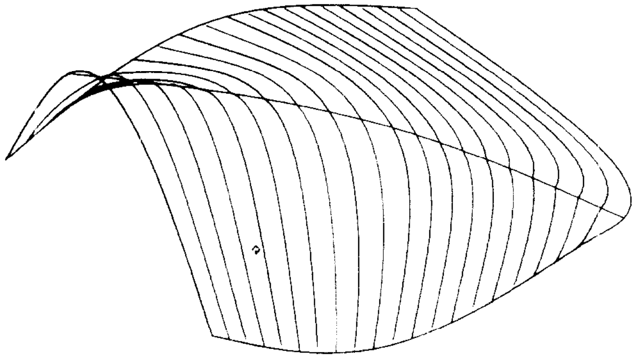
In practice, one usually aims for surfaces which are smooth at least to the extent that the surface normal is well defined at each point. (We shall ignore the delicate issues involved in treating the singularities arising at nonordinary points, which may or may not be removable.) For surfaces constituted out of assemblies of bicubic patches interpolated through rectangular data the simplest condition ensuring this C^1 condition is that the triad of vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_{12}$ assigned at each patch corner (i.e., data point) should be the same for adjacent patches. For then the transverse tangents on either side of a common boundary will each be the same linear combination of the same vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_{12}$ at the end-points, hence the surface normal will be continuous as one crosses over the boundary. (There are many variations on this condition; see [2, Chap. 7].) If cubic splines are used to define the patch boundaries of such a composite surface, then the tangents relative to the global spline parameters will be determined; let us suppose that the twists are also supplied relative to these global parameters (not necessarily by the method described above). Then in order to apply the Hermite form (1.1) (which assumes a $[0, 1]$ parameter interval) to each patch separately, it is necessary to scale the global vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_{12}$ by factors equal to the appropriate global parameter intervals that have been chosen for the splines, before inserting them into the matrix Q (see [2, Sect. 7.2.4]). There now arises a problem. If the global spline parameter intervals between successive boundary curves are different, the scaling factors appearing in (2.1) will be different for adjacent patches and the transverse tangents will now be different linear combinations of the original global $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_{12}$ in adjacent patches; C^1 smoothness can no longer be guaranteed. This happens, for example, when chord-length parameterization (the preferred method in the univariate case) is used for the splines. On the other hand, if uniform parameter intervals are used for these splines, C^1 continuity is achieved only at the risk of loops or unwanted bulges appearing between points on the spline that are too close together. Now it has been observed in practice that the former option does in most cases actually yield acceptable smooth surfaces. We shall attempt to explain this in terms of the results of Section 3. Thus suppose the patch boundary curves through the data points are splines with any chosen parameter intervals between the knots. Then global vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_{11}$, $\mathbf{r}_{22}$ and $\hat{n}$ are available and continuous at each data point. (The subscripts here denote differentiation with respect to the global spline parameters, of course.) Let us suppose also that some global twist vector is assigned at each data point. Then, each of the four resulting bicubic patches sharing some point as a common corner gives rise to a Weingarten matrix (see Section 3), which in local patch parameters assumes the form

$$W_{\text{local}}^{(i,j)} = -\begin{bmatrix} h_i^2 L & h_i k_j M \\ h_i k_j M & k_j^2 N \end{bmatrix} \begin{bmatrix} h_i^2 E & h_i k_j F \\ h_i k_j F & k_j^2 G \end{bmatrix}^{-1}, \qquad i, j = 1, 2,$$

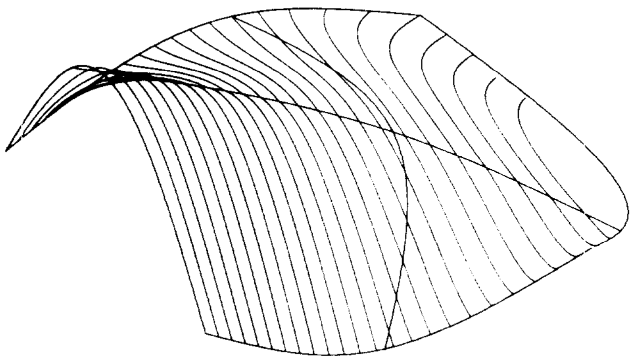
where h_1, h_2 (k_1, k_2) are consecutive parameter intervals as one passes along both splines through the data point in question, and where $E, \ldots, N$ are defined in terms of these global parameters. Rearranging, we have for all i, j

$$\begin{aligned} W_{\text{local}}^{(i,j)} &= -\begin{bmatrix} h_i & 0 \\ 0 & k_j \end{bmatrix} \begin{bmatrix} L & M \\ M & N \end{bmatrix} \begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{bmatrix} h_i & 0 \\ 0 & k_j \end{bmatrix}^{-1} \\ &= \begin{bmatrix} h_i & 0 \\ 0 & k_j \end{bmatrix} W_{\text{global}} \begin{bmatrix} h_i & 0 \\ 0 & k_j \end{bmatrix}^{-1}. \end{aligned}$$

Thus the matrices $W_{\text{local}}^{(i,j)}$ are all similar to the matrix representing the global Weingarten map and hence are all similar to each other at the common corner. Consequently the patches abutting this point have local expansions (3.3) relative to the same local axes, which are identical up to the terms of second degree. That is, the *actual* surfaces differ only in third (and higher) degree terms of their (ξ, η)-expansions (3.3). Consequently, there will be a neighborhood about each data point in which the departure from continuity of the surface normal along patch boundaries will be below the resolution of the intersection routines or plotting device tolerances. The size of these neighborhoods will of course depend upon these tolerances and on the higher order terms in the expansion (3.3). The closer analysis entailed by these considerations, which should also take advantage of the fact that the surfaces are bicubics, will not be entered into here. Figure 5 illustrates this effect on a four patch surface whose boundaries are cubic splines with natural end conditions and chord



a

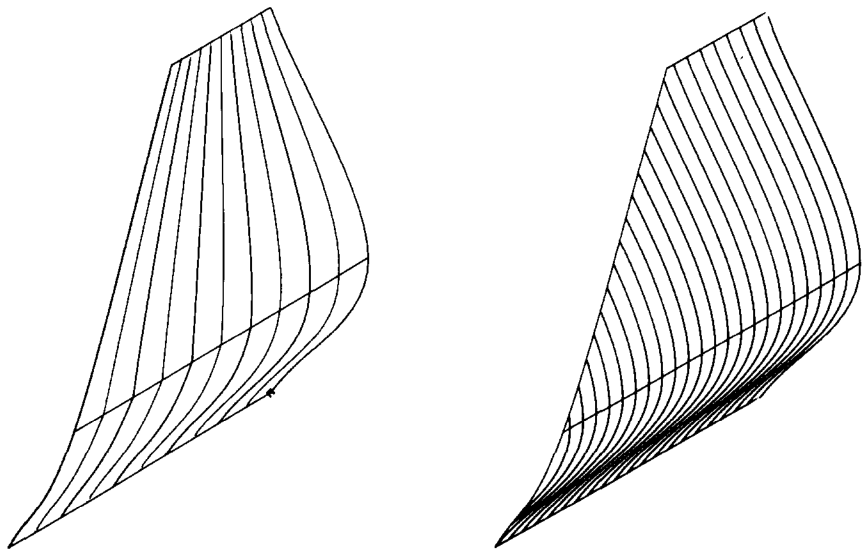


b

FIG. 5 (a) A four patch surface with chord length parametrization used for boundary splines. Vertical curves are constant parameter lines. Note departure from C^1 smoothness along central boundary. (b) Same surface as in (a) with planar sections shown. Note apparent smoothness along central boundary.

length parameterization. The chord lengths were deliberately chosen to vary significantly, at least in one direction, along the surface. Figure 5(a) shows a plot of one set of parameter lines (which does not involve any intersection routines). The increasing departure from C^1 smoothness of these curves as one moves along the central spline between data points is clearly visible. Figure 5(b) shows a family of planar sections; the rate of departure from C^1 smoothness along the same common boundary is much less marked, as expected.

It will be noticed that the foregoing argument requires only the continuity of the entries in the (global) Weingarten matrix; in particular the twist vector participates only through its normal component. This indicates that the approximation to C^1 smoothness just described could be achieved for patch assemblies if, in addition to the continuity of the univariate quantities, one imposes continuity only upon M, the normal component of the twist, leaving the tangential components to be chosen independently in each patch. Since these components control the local disposition of the parameter lines (cf. (1.5), (1.6)) this could be advantageous in certain applications. Figure 6 illustrates this for a bent panel-like surface comprising two bicubic patches. The vertical boundaries are splines with unit parameter intervals between the knots and clamped end conditions. The normal components of the twists have



a b

FIG. 6. (a) A two patch surface with the normal component only, of the twist, continuous along boundary splines. Vertical curves are constant parameter lines. Note departure from C^1 smoothness along central boundary. (b) Same surface as in (a) with planar sections shown. Note apparent smoothness along central boundary.

been chosen according to the method described in Section 1 with values of K arbitrarily assigned. The tangential twist components have been chosen independently in each patch deliberately to exaggerate the effect. The departure from C^1 smoothness is again clearly visible along the common boundary in the parameter-line plot shown in Fig. 6(a). The planar sections shown in Fig. 6(b) appear quite smooth, however.

It is relevant at this point to remark that if (1.4) is invoked to calculate M then either global or local patch parameters may be used, since K is dimensionless. That is to say, this method of extracting the normal twist component commutes with operation of changing parameters. This may not be true for the tangential twist components; for example, it will in general be false if finite differences are used to estimate the second derivatives on the right-hand sides of (1.5), (1.6) in the case of a nonuniform parameter spacing between successive boundary curves.

3. MATHEMATICAL APPENDIX

A parametric surface is locally the image of a function of the form

$$\mathbf{r}: U \to \mathbb{R}^3,$$

where U (the "parameter space") is an open subset of $\mathbb{R}^2$ ($\mathbb{R}^n$ denoting Euclidean n-space). Thus we may write

$$\mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v))$$

and differentiability properties assumed for $\mathbf{r}$ will be inherited by the functions

$x(u, v)$, $y(u, v)$ and $z(u, v)$. If

$$\phi : V \to U$$

is a one-to-one function (where V is open in $\mathbb{R}^2$) having differentiability properties similar to those of $\mathbf{r}$, then the image of the composition

$$\mathbf{r} \circ \phi : V \to \mathbb{R}^3$$

is the same surface as before (i.e., the same set of points in $\mathbb{R}^3$) but with a new (local) parameter space (V) and a new "parameterization" $\mathbf{r} \circ \phi$. We shall be interested in certain properties of surfaces which are *locally invariant*; i.e., properties that are independent of (local) reparameterizations. (Among these one hopes to find properties that reflect the genuine geometrical nature of the surface, not merely effects which depend upon a particular (u, v)-parameterization.)

For a point p on the surface, we denote by T_p the tangent plane at p, and by $\hat{n}_p$ the unit surface normal at p. (Except for one example, we consider only *ordinary* points; i.e., those at which these entities are well defined. Moreover, when p is not in doubt we shall frequently drop the subscript.) These notions are clearly invariant. A linear map (the *Weingarten map*)

$$W_p : T_p \to T_p$$

is defined for a tangent $\mathbf{X} \in T_p$ by

$$\begin{aligned} W_p(\mathbf{X}) &= (\mathbf{X} \cdot \nabla)\hat{n}_p \\ &= x_1 \frac{\partial \hat{n}_p}{\partial x} + x_2 \frac{\partial \hat{n}_p}{\partial y} + x_3 \frac{\partial \hat{n}_p}{\partial z}, \end{aligned} \tag{3.1}$$

where $\mathbf{X} = (x_1, x_2, x_3) \in \mathbb{R}^3$. (The operator $\mathbf{X} \cdot \nabla$, for any vector $\mathbf{X}$, can be interpreted as "differentiation along $\mathbf{X}$.") This linear map is an invariant since it is defined independently of the parameterization. Consequently those functions of linear maps that are basis independent yield surface invariants when applied to the Weingarten map. Thus, for example, the determinant, trace, eigenvalues and eigenvectors of W_p are all (local) surface invariants at p, called, respectively, Gaussian curvature, mean curvature, principal curvatures, principal directions. (W_p can be shown to be self-adjoint so that its eigenvalues are real.) We can derive formula (1.3) for the Gaussian curvature K immediately from the definition. Having chosen some (u, v) parameterization, it is quickly seen that

$$\begin{aligned} W(\mathbf{r}_i) &= \hat{n}_i \\ &= \alpha_{i1}\mathbf{r}_1 + \alpha_{i2}\mathbf{r}_2, \qquad i = 1, 2, \end{aligned}$$

for some matrix (α_{ij}), $j = 1, 2$. Then

$$\begin{aligned} \hat{n}_1 \wedge \hat{n}_2 &= (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})\mathbf{r}_1 \wedge \mathbf{r}_2 \\ &= (\det W)\mathbf{n} \\ &= K\mathbf{n}, \end{aligned}$$

where the wedge denotes ordinary cross product in $\mathbb{R}^3$. Thus

$$K = \frac{\hat{n}_1 \wedge \hat{n}_2 \cdot \mathbf{n}}{\mathbf{n}^2}.$$

The classical "equations of Weingarten" give an explicit expression for the matrix (α_{ij}). With notation as before these equations assert that relative to the basis $\mathbf{r}_1$, $\mathbf{r}_2$ of T, the Weingarten map assumes the matrix from

$$W = -\Omega g^{-1}, \tag{3.2}$$

where

$$g = \begin{bmatrix} E & F \\ F & G \end{bmatrix} \quad \text{and} \quad \Omega = \begin{bmatrix} L & M \\ M & N \end{bmatrix}$$

(see [4, Eq. 9.15, Chap. III Sect. 9] and recall that $W(\mathbf{r}_i) = \hat{n}_i$, $i = 1, 2$).

From this one obtains the usual formula (1.2) for the Gaussian curvature:

$$\begin{aligned} K &= \det W \\ &= (-1)^2 \det \Omega / \det g \\ &= LN - M^2/H^2 \\ &= \frac{(\hat{n}\cdot\mathbf{r}_{11})(\hat{n}\cdot\mathbf{r}_{22}) - (\hat{n}\cdot\mathbf{r}_{12})^2}{\mathbf{r}_1^2\mathbf{r}_2^2 - (\mathbf{r}_1\cdot\mathbf{r}_2)^2}. \end{aligned}$$

From this formula one notes the interesting fact that the vanishing of the twist at a point produces in general a *nonzero* curvature at the point. This runs counter to the commonly expressed belief that a surface should be "flat" at such points. The effects of this nonzero curvature at the point in question (produced, for instance, at patch corners when the twists are taken to be zero) would not necessarily be apparent in a plot since a large number of densely packed sections may be required to show it up. Nevertheless the effects of this nonzero curvature will propagate through the surface and may be observed to produce unwanted rippling effects elsewhere (cf. [1]). On the other hand, a point at which the surface is genuinely flat (i.e., $K = 0$) will in general actually require a *nonzero* value for the (normal component of the) twist (cf. Fig. 1).

Although a knowledge of the Gaussian curvature at a point on a surface is useful in specifying the normal component of the twist in the manner described in Section 1, it is not particularly revealing in itself about the nature of the surface in the vicinity of the point. However once M has been found (or assigned) so that all the entries in the Weingarten matrix (3.2) may be assumed known, then there is enough information to actually determine the surface up to second order in a neighborhood of the point. To show this, we return to the Weingarten map W_p defined at some point p (3.1) and suppose first that W_p has distinct eigenvalues. It is a standard elementary result that the corresponding eigenvectors in T_p (i.e., the principal directions) are orthogonal, and these directions, together with the surface normal direction at p, define an orthogonal triad in $\mathbb{R}^3$ with origin at p. Referred to these axes, the surface may now be locally (re-)parameterized by the map sending (ξ, η) to

(ξ, η, ζ) where ξ and η are measured along the principal directions and ζ is the perpendicular distance in the direction $\hat{n}_p$ of the corresponding point on the surface. (A unique choice of surface point is possible in a neighborhood of an ordinary point p.) Writing $\mathbf{r}(\xi, \eta) = (\xi, \eta, \zeta)$ we have

$$\begin{aligned}\zeta &= \boldsymbol{r}(\xi, \eta)\cdot\hat{n}_p\\ &= \tfrac{1}{2}(L\xi^2 + 2M\xi\eta + N\eta^2) + O(\xi^3, \eta^3),\end{aligned}$$

on applying Taylor's theorem. (Here L, M, N are of course calculated relative to (ξ, η).) It follows that in this parameterization the Weingarten matrix has the form

$$-\begin{bmatrix} L & M \\ M & N \end{bmatrix}.$$

Now it is easy to show that an orthogonal transformation of the ambient Euclidean 3-space leaves the Weingarten map unchanged, and since the (ξ, η)-axes were chosen to lie along eigenvectors of the Weingarten map arising when the original (x, y, z)-axes were used, this remains the case when (ξ, η, ζ)-coordinates are used in the ambient space. Thus the Weingarten map is diagonal in (ξ, η)-coordinates and consequently $M = 0$. The local equation of the surface then becomes

$$\zeta = \tfrac{1}{2}(L\xi^2 + N\eta^2) + O(\xi^3, \eta^3), \tag{3.3}$$

where L, N are (up to a constant sign) the eigenvalues of W_p. Thus the Weingarten map at p canonically determines the surface in a neighborhood of p (up to second order) since all the ingredients (up to second order) in the above formula depend only upon similarity invariants of W_p (i.e., principal directions, principal curvatures).

A similar argument applies in case W_p is a multiple of the identity, if (ξ, η)-axes are chosen arbitrarily in T_p. In this case an analogous formula (1.3) holds with $L = N$, but of course the axes are not canonically determined by W_p.

One notes that in this parameterization, the twist vanishes at p.

4. CONCLUSIONS

We have sought to identify the geometrical issues involved in the choice of twist vectors in the modeling of surfaces by relating them to local invariants. The Gaussian curvature emerges as a significant invariant that is intrinsic to the surface (hence in principle calculable independently of parameterizations) but which can be related to the normal twist component. Thus the twist can be chosen in such a way that the modeled surface matches the original surface more closely to the extent that the Gaussian curvature is correct at the data points. The normal twist component emerges, through its role in the Weingarten matrix, as the primary bivariate determinant of the local topology of the surface, and consequently care must be exercised in its choice. The tangential twist components can be treated in a more cavalier fashion, to particularly good effect in applications not requiring (or admitting) full C^1 smoothness.

ACKNOWLEDGMENTS

Work for this paper was done on behalf of Pressed Steel Fisher Ltd., Oxford, U.K. Thanks are owed to Peter Charrot (PSF) and Terry Stacey (Oxford University

Computing Laboratory and PSF) for many helpful conversations and to them and Alan Bale, Tony Coy and Ken Maynard (all of PSF) for programming assistance. In particular, Dr. Charrot was responsible for a suite of programs implementing some of the methods described here which produced Fig. 5. A large debt of gratitude is owed also to Douglas Brand of the PSF CAD-group for his generous support and encouragement during the course of this work.

Thanks are owed also to the referee for several pertinent comments.

REFERENCES

1. R. E. Barnhill, J. H. Brown, and I. M. Klucewicz, A new twist in computer aided geometric design, *Computer Graphics and Image Processing* **8**, 1978, 78–91.
2. I. D. Faux and M. J. Pratt, *Computational Geometry for Design and Manufacture*, Wiley, New York, 1979.
3. N. J. Hicks, *Notes on Differential Geometry*, Van Nostrand, Princeton, 1965.
4. T. J. Willmore, *An Introduction to Differential Geometry*, Oxford Univ. Press, New York, 1959.
5. G. Joos, *Theoretical Physics*, Blackie, London, 1953.