Abstract

This paper addresses the long-standing problem of the unavoidable gaps that arise when expressing the intersection of two NURBS surfaces using conventional trimmed-NURBS representation. The solution converts each trimmed NURBS into an untrimmed T-Spline, and then merges the untrimmed T-Splines into a single, watertight model. The solution enables watertight fillets of NURBS models, as well as arbitrary feature curves that do not have to follow iso-parameter curves. The resulting T-Spline representation can be exported without error as a collection of NURBS surfaces.

CR Categories: I.3.5 [Computer Graphics]: Computational geometry and object modeling—Curve, surface, solid, and object representations

Keywords: Surface intersection, Booleans, NURBS, T-Splines

1 Introduction

The trimmed-NURBS modeling paradigm suffers from a serious fundamental flaw: parametric trimming curves are mathematically incapable of fulfilling their primary role, which is to represent the curve of intersection between two NURBS surfaces. Consequently, trimmed-NURBS models are not mathematically watertight, as illustrated in Figure 1 in which trimming curves are used to express the intersection between the body and spout of the Utah teapot model. We use the term “watertight” to connote no unwanted gaps or holes. That is, the surface is a gap-free 2-manifold in the neighborhood of intersection curves.

This paper presents a two-step algorithm for representing the Boolean combination of two NURBS objects as a single watertight T-Spline. In the first step, each trimmed NURBS is converted into a T-Spline without trimming curves, as illustrated in Figure 2.a. In the second step, each pair of untrimmed T-Splines is merged along their intersection curve into a gap-free T-Spline model, as illustrated in Figure 2.b. The resulting model, shown in Figure 2.c is $C^2$, except at the $C^0$ crease along the intersection curve. This T-Spline model facilitates the creation of watertight fillets. The model in Figure 2.d contains a $C^2$ gap-free fillet between the body and spout.

Section 2 reviews the history and significance of the problem this paper addresses, and reviews prior literature. Section 3 presents an algorithm for converting a trimmed-NURBS into an untrimmed T-Spline. Section 4 explains how to merge two NURBS or T-Spline surfaces with mis-matched parametrizations. Section 5 details how the algorithms presented in Sections 3 and 4 work together to create watertight trimmed-NURBS models, and examines the approximation error. This section also discusses the creation of gap-free, $C^2$ fillets and the placement of feature lines on a T-Spline surface that are not aligned with iso-parameter curves. Section 6 summarizes.

2 Background

The fact that gaps are unavoidable in conventional trimmed-NURBS mathematical models can be shown as follows. A trimming curve is typically a degree-three NURBS curve defined in the parameter domain of a NURBS surface. The image of such a trimming curve on a bicubic patch (i.e., the curve on the bicubic patch in
that the trimming curve maps to) is degree ≤ 18 and algebraic genus zero. However, a generic intersection curve of two bicubic surfaces is degree 324 in $\mathbb{R}^3$ [Sederberg et al. 1984] and algebraic genus 433 [Katz and Sederberg 1988]. Hence, intersection curves can only be approximated by parametric trimming curves. The existence of these gaps in trimmed NURBS models seems innocuous and easy to address, but in fact it is one of the most serious impediments to interoperability between CAD, CAM and CAE systems [Kasik et al. 2005]. Software for analyzing physical properties such as volume, stress and strain, heat transfer, or lift-to-drag ratio will not work properly if the model contains unresolved gaps. Since 3D modeling, manufacturing and analysis software does not tolerate gaps, humans often need to intervene to close the gaps. This painstaking process has been reported to require several days for a large 3D model such as an airplane [Farouki 1999] and was once estimated to cost the US automotive industry over $600 million annually in lost productivity [NIS 1999]. At a workshop of academic researchers and CAD industry leaders [Farouki 1999], the existence of gaps in trimmed-NURBS models was singled out as the single most pressing unresolved problem in the field of CAD.

Prior Art
Several solutions to the gap problem have been put forward, but none address the problem adequately. The best solution from a theoretical standpoint is to use the precise representation for trimming curves, which is an implicit (not parametric) equation of the form $f(s, t) = 0$. In the case of two intersecting bicubic patches, $f(s, t)$ is a polynomial of bi-degree $54 \times 54$. [Krishnan and Manocha 1996] presents a solution to the surface intersection problem based on such a representation, and [Krishnan et al. 2001] describes a solid modeling system based on this approach, using exact arithmetic. Unfortunately, exact arithmetic can be very expensive and the method has not been adopted by the CAD industry. In applications for which a tessellation of the surfaces suffices, gaps can easily be filled with a triangle strip or avoided altogether by careful coordination while tessellating adjoining trimmed surfaces [Kumar 1996; Moreton 2001]. However, once a NURBS model has been reduced to a $C^0$ tessellation, it loses its character as a smooth surface and operations such as offsetting become impossible.

[Song et al. 2004] and [Farouki et al. 2004] describe methods for creating a non-tessellated, watertight approximation of two or more intersecting NURBS surfaces. Each method produces a set of piecewise $C^0$ (but approximately $C^1$) Bézier patches, although if patches adjacent to an intersection curve are edited, the surfaces become discontinuous. Our technique produces a watertight $C^2$ surface defined using a single T-Spline control grid, so the surface remains $C^2$ if the control points are moved. [Song et al. 2004] requires the solution of a system of linear equations that under some conditions can produce huge approximation errors. Our method is more amenable to creating fillets than [Song et al. 2004] and [Farouki et al. 2004].

Kristjansson et al. 2001] takes as input Loop subdivision surfaces, although extension to other types of subdivision surfaces is possible. [Kristjansson et al. 2001] produces a $C^2$ watertight subdivision surface defined by a multi-resolution control grid; the surface remains $C^2$ if the control grid is edited. The goal of [Kristjansson et al. 2001] is an efficient algorithm suitable for animation, but not necessarily for CAD.

A pertinent prior art to Section 3 is [Litke et al. 2001], which describes a process of converting a trimmed subdivision surface into an untrimmed subdivision surface such that each trimming curve on the trimmed surface becomes a boundary curve on the untrimmed surface. [Litke et al. 2001] uses an enhanced Loop surface (triangle based), whereas our algorithm is based on tensor-product NURBS surfaces. Both algorithms must perturb the surface in the neighborhood of each trimming curve, but are capable of confining the perturbation to an arbitrarily small magnitude and narrow neighborhood. [Litke et al. 2001] uses trimming curves that lie in world space, thus permitting the true intersection curve to serve as the trimming curve, whereas we use conventional parametric trimming curves, which can only approximate true intersection curves.

A key disadvantage of subdivision surfaces for use in CAD is their incompatibility with NURBS. Billions of dollars have been invested in NURBS software and models, and there is tremendous economic pressure against abandoning the NURBS paradigm. Furthermore, NURBS do have some advantages over subdivision surfaces. For example, numerous versions of subdivision surfaces exist, with no current industry standard, and many capabilities of subdivision surfaces involve special refinement rules. Also, subdivision surfaces are limit surfaces, involving infinite sequences of patches. Although there are efficient ways to evaluate subdivision surfaces [Stam 1998], the infinite number of patches is more difficult to deal with than a finite number of NURBS patches, especially when doing data file exchange. One approach used in commercial CAD software to manage the NURBS gap problem is to use a procedural definition of intersection curves, which keeps track of which surfaces intersect. Intersections can then be approximated, on demand, to any desired tolerance. This approach complicates subsequent tasks such as offsetting or filleting that require an explicit representation of the intersection curve. Furthermore, if a procedural definition of the same intersection is used by two different programs, it is possible to arrive at different results. This has resulted in incompatibilities between NURBS representations by different CAD, CAM, and CAE software applications, and the growth of an entire software industry around translating, fixing and healing 3D models and surfaces. The surface intersection problem has been very thoroughly researched. A sampling of the vast literature can be found in [Patel and Chang 2002; Song et al. 2004]. The algorithms described in this paper assume the existence of a robust surface intersection algorithm that can represent an intersection curve using trimming curves to within a prescribed tolerance, such as in [Kristjansson and Manocha 1997]. Such capability is now standard in most commercial geometric modeling programs.

The topic of fillets has likewise been widely researched. See [Song and Wang 2007] for a list of references, and a solution to the fillet problem that provides $G^2$ continuity. Most commercial software approximates fillets of free-form surfaces as NURBS surfaces that lie on the base surfaces, with approximate $G^2$ continuity. The advantage of the fillet solution presented in this paper is that it is part of a unified geometric framework. An entire geometric model, including fillets, can be represented as a single watertight T-Spline.

3 Trimmed-NURBS to Untrimmed T-Splines
This section presents a method for converting a bicubic NURBS surface with trimming curve into an approximately equivalent T-Spline with no trimming curve. The approximation error can be made arbitrarily small, and the perturbation can be confined to an arbitrarily narrow neighborhood of the trimming curve. We describe the algorithm using the example in Figure 3. Figure 4.a diagrams the trimming curve $C$ in the parameter domain of the NURBS surface. The grid lines are knot lines for the NURBS surface. Points in the domain that correspond to NURBS control points are highlighted in red.

**Trimmed-NURBS to Untrimmed T-Splines Conversion**
Step 1. Form an axis-aligned polygon $A$ (i.e., a polygon whose edges are parallel to one of the two parameter directions) that en-
closes the trimming curve, as illustrated in Figure 4.b.

Step 2. At each vertex of \( A \) that does not lie on a red point, perform a T-Spline control point insertion as described in [Sederberg et al. 2004]. In this example, the control point insertions will occur at the five red points lying on the black line in Figure 4.c. (The insertion operation adds two additional control points to the left of the black corner, and two beneath it, as shown in Figure 4.d.)

Step 3. Remove the portion of the control mesh that lies on the interior of \( A \) and replace it with the mesh topology illustrated in Figure 4.d. Note that each convex corner in \( A \) introduces a valence three control point in the modified control grid, and each concave corner in \( A \) creates a valence five control point. Assign a knot interval of zero to all edges of the control grid that connect a blue control point to a green control point, as shown in Figure 5.a. These zeros create a Bézier end condition for this boundary curve. Assign a small knot interval \( \alpha \) to all edges connecting the outer layer of green control points to the inner layer of green control point. A good choice for \( \alpha \) is the average of all parameter distances between the vertices on \( A \) and the trimming curve.

Step 4. Leave all red control points in Figure 4.d in their initial location. The blue control points in Figure 4.d define a NURBS curve that approximates the image of \( C \). The positions of those control points are chosen to minimize the orthogonal distance between the NURBS curve and the image of \( C \), using an algorithm such as in [Wang et al. 2006]. Likewise, the positions of the green control points are chosen to minimize the orthogonal distance between the T-Spline surface and the interior of the trimmed NURBS surface. The resulting T-Spline and its control grid are shown in Figure 3.b.

This procedure introduces some perturbation error, the magnitude of which in this example is 0.001 times the width of the model. The domain of the perturbed region lies within the support of the green and blue control points. Figure 5.a illustrates the perturbation region in yellow. For a fixed axis-aligned polygon, we can make the perturbation region on the exterior of the polygon arbitrarily narrow by performing a local T-Spline refinement, as illustrated in Figure 5.b. Likewise, we can make the distance between \( A \) and \( C \) arbitrarily small by finding an axis-aligned polygon that approximates \( C \) to within a tolerance \( \epsilon \). Clearly, there are countless such axis-aligned polygons, and numerous possible algorithms for finding such polygons. We now present one such algorithm.

### 3.1 Finding an Axis-Aligned Polygon

The algorithm, illustrated by the example in Figure 6, has the flavor of a curve rasterization in which the pixels are cells of a quadtree. Related applications of quadtrees are reported in [Hunter and Stei-]
Table of Contents
1 Introduction
2 Related Work
3 Procedure
4 Merging Using NU-NURBS
5 Conclusion
Acknowledgments
References
An improved algorithm for computing the green control points is a problem calling for future research. To create our example figures, we chose a traditional method that appeared simplest to implement. We begin by obtaining a set of sample points on the region of the trimmed surface that will be perturbed and then specify an initial position for the green control points. The following process is then repeated: Assign each sample point a parameter pair on the untrimmed T-Spline, then solve for the green control points that minimize the least squares error based on these parameter assignments. It is known that this algorithm converges only linearly [Bjorck 1996], and indeed it can take several tens of seconds to obtain good results if the initial positions of the green control points are not chosen wisely. The algorithm has also been observed to converge to a local min.

One possible solution is to extend to surfaces the curve-fitting algorithm in [Wang et al. 2006], something the authors of that paper are working on. Another line of research is to study whether the quasi-interpolation methods used in [Litke et al. 2001] can be adapted to T-Splines.

4 Merging Using NU-NURBS
After two intersecting surfaces are converted into untrimmed T-Splines using the method in Section 3, the final step is to merge those two T-Splines into a single, gap-free T-Spline. A basic algorithm for merging two T-Splines is presented in [Sederberg et al. 2003]. However, that algorithm gives poor results if the two surfaces to be merged do not have consistent parametrizations, as illustrated in Figure 8. The first step in the merge algorithm in [Sederberg et al. 2003] is to insert knots such that the two surfaces have the same set of knot intervals, as shown in Figure 8.b. (This might also require that all knot intervals on one surface be scaled so that their sum matches the sum of the knot intervals on the other surface). The final step is to connect the two control grids, as shown in Figure 8.c. However, if the adjoining boundary curves are not parametrized similarly, the isoparameter curves will experience an abrupt bend, imparting a kink in the resulting surface. For example, Figure 9 shows a hand and arm modeled as separate NURBS, whose parametrizations do not align. Figure 9.b shows the result of merging them using the algorithm described in [Sederberg et al. 2003]. Unfortunately, most pairs of untrimmed T-Splines generated using the method in Section 3 have this problem.

The problem is related to the fact that the refinements shown in Figure 8.b must honor a restriction that is placed on the knot intervals in a T-Spline: the sum of knot intervals on one edge of a face on the control grid must equal the sum of knot intervals on the opposing edge in the face. Better results could be obtained by, instead of refining each surface as in Figure 8.b, we refine the two surfaces so that their knot lines align, as shown in Figure 8.d. However, the resulting knot interval configuration violates the definition of a T-Spline. Previous methods for dealing with such knot intervals [Sederberg et al. 1998; Möller et al. 2006] devise variations on Catmull-Clark refinement in which faces of the control grid map to an infinite sequence of bicubic patches which, like extraordinary points in a Catmull-Clark surface, are $G^2$. The infinite sequence of patches violates a key objective of this paper, which is to be ex-
portable using a finite number of tensor-product patches.

To address this problem, we introduce a generalization of tensor-product B-Spline surfaces that supports the knot interval configuration in Figure 10.b, that is $C^2$, and that yields one tensor-product patch per face of the control grid. Since the knot intervals change and hence are not “uniform,” we will refer to this surface as a non-uniform NURBS surface, or NU-NURBS (spoken “new NURBS”).

For a fixed value of $t = \tau$,

$$P(s) = \sum_{i=1}^{m} P_i(\tau) B_i(s)$$

defines a cubic B-Spline curve that lies on the bicubic B-Spline surface, and is the iso-parameter curve for $t = \tau$. If we let $\tau$ vary, the resulting family of iso-parameter curves sweeps out the B-Spline surface. Figure 11.b shows four such iso-parameter curves for various values of $\tau$.

Since most readers will be more familiar with B-Splines defined using knot vectors rather than with knot intervals, we note that it is straightforward to convert between the two representations. Define $\bar{e}_{-2} = 0, \bar{e}_{i+1} = \bar{e}_i + e_i, i = -2, \ldots, 5$ (where $e_i$ and $e_0$ are not shown in the figure). Then the knot vector for the B-spline curves $P_i(t)$ is $\{\bar{e}_{-2}, \bar{e}_{-1}, \ldots, \bar{e}_6\}$. Likewise, the knot vector for each of the iso-parameter curves in Figure 11.b is $\{\bar{d}_{-2}, \bar{d}_{-1}, \ldots, \bar{d}_7\}$ where $\bar{d}_{-2} = 0, \bar{d}_{i+1} = \bar{d}_i + d_i, i = -2, \ldots, 6$.

We now modify that description of a B-Spline surface to permit a knot interval arrangement as in Figure 12.a in which the $d_{ij}$ can be any non-negative number. The basic idea is to treat the knot intervals themselves as cubic spline functions, which in turn control the basis functions $B_i(s)$ in (2).
This NU-NURBS formulation has the following properties:

1. It specializes to NURBS in the case where the knot interval configuration is identical to that in Figure 10.a.

2. This NU-NURBS is \( C^2 \) in \( s \), since each iso-parameter curve \( \mathbf{P}(s) \) is a cubic spline curve whose knot intervals are constant for a fixed value of \( t \). Since the knot interval functions are \( C^2 \) splines in \( t \), the NU-NURBS is also \( C^2 \) in \( t \).

3. The cost of evaluating this NU-NURBS is comparable to the cost of evaluating a bicubic NURBS surface, the only difference lies in evaluating the knot interval spline functions \( d_i(t) \).

4. Although this surface formulation is new and hence not directly supported in existing commercial software, it can be exactly represented—and exported—as a set of rational Bézier patches. In our T-Spline model, the body is defined using 90 control points. We have devised other, lower-degree versions of NU-NURBS, all that produce one patch per face, including a version that is \( C^2 \) in \( t \) and \( C^1 \) in \( s \) with patches that are degree \( 3 \times 6 \), and one that is \( C^2 \) in \( t \) and \( C^2 \) in \( s \) with patches that are degree \( 4 \times 9 \). We have also devised a version of NU-NURBS that permits arbitrary non-negative knot intervals in both parameter directions. These variations are being recorded in a separate paper [Sederberg et al.].

## 5 Examples

This section examines the behavior of the algorithms presented in Sections 3 and 4 when they combine to represent two intersecting trimmed-NURBS surfaces as a single watertight T-Spline. It also shows how fillets are supported in this representation, along with arbitrary feature lines.

As reviewed in Section 2, the problem of computing intersection curves is very well studied, and algorithms for computing the intersection of two NURBS surfaces \( P_1 \) and \( P_2 \) are standard in most geometric modeling programs. These algorithms can compute trimming curves \( C_1 \) and \( C_2 \) in the parameter domains of \( P_1 \) and \( P_2 \), along with an approximation \( C \) of the intersection curve in \( \mathbb{R}^3 \), to within a prescribed tolerance. This paper assumes that \( C_1 \), \( C_2 \), and \( C \) have been computed using an existing algorithm, and that geometric and topological accuracy of these curves is deemed acceptable.

### Perturbation Error

Figure 13 shows the trimmed teapot body being converted into untrimmed T-Splines with different degrees of precision. The NURBS model of the teapot used throughout this paper is actually a \( C^2 \) NURBS model based on the original \( C^1 \) Bézier model. In our NURBS model, the body is defined using 90 control points.

Figure 14 shows the error distribution in both the body and the spout. This figure shows how the perturbation domain decreases as \( \epsilon \) decreases.

By adjusting \( \epsilon \), the perturbation magnitude and extant can be held below a specified tolerance.

The perturbation errors reported in the captions in Figure 13 are relative to a teapot that is one unit wide. Hence, the error in Figure 13.a is about one tenth of one percent of the width of the teapot.

### Fillets

Figure 16.a shows the knot intervals adjacent to the intersection curve. Immediately after the merge is completed, \( k = 0 \). This creates a triple knot at the intersection curve and forces the T-Spline to be \( C^0 \) along the intersection curve. If the value of \( k \) is changed to a small positive value, the \( C^0 \) crease along the intersection curve is changed into a \( C^2 \) fillet whose radius increases with \( k \).

### Feature Lines

Feature lines in NURBS models must follow iso-parameter curves. T-Splines allow for sharp features along portions of iso-parameter
and can be exported without translation error as a finite collection of NURBS patches, making them compatible with CAD industry standards. The models are editable in that control points can be adjusted and the surface will remain $C^2$. These results can help to streamline the CAD modeling–analysis pipeline. In addition, these tools introduce new design workflows into the styling and CAD industries, allowing NURBS modelers to continue to style their models even after Booleans have been performed.

The paper also presents an enhanced merging capability involving an augmentation of the definition of T-Splines to support different knot intervals on opposite sides of the same control polygon face. This enables the merging of two NURBS or T-Splines surfaces whose mating curves are parametrized differently, a case not handled well in [Sederberg et al. 2003].

Our process for merging two trimmed NURBS surfaces into an untrimmed T-Spline involves a perturbation of the original surfaces. The perturbations can be limited to an arbitrarily narrow strip. Tighter tolerances demand more control points, making the resulting T-Spline more difficult to edit, although if the designer’s intent is to ultimately edit the resulting T-Spline, a high initial tolerance may not be as crucial.

This paper invites future research on several fronts. Section 3 discusses possible lines of research for finding an efficient algorithm for computing the green control points, the problem we view as most pressing. Also called for is a rigorous analysis of approximation error. What is the relationship between $\epsilon$, positional error, and normal-vector error? What is the convergence rate?

The paper focuses on two intersecting surfaces. Details of how to handle three or more intersecting surfaces are not presented and invite further study. The existing algorithms should extend readily to handle most cases where all intersections are to be computed simultaneously. The case where an untrimmed T-Spline that represents two intersecting surfaces is later intersected by a third NURBS or T-Spline is more challenging because it can involve the intersection of NU-NURBS or of faces next to extraordinary points. It is also not clear how the error might propagate upon repeated such intersections.

The paper only addresses non-singular intersection curves. Full treatment of intersection curves that self-intersect is another topic of further study.

7 Acknowledgements

Nicholas North and Adam Helps provided invaluable assistance with implementation and figures.
References


