Abstract

We present a solution to the general problem of characterizing shadows in scenes involving a uniform polygonal area emitter and a polygonal occluder in arbitrary position by manifesting shadow irradiance as a spline function. Studying generalized prism-like constructions generated by the emitter and the occluder in a four-dimensional (shadow) space reveals a simpler intrinsic structure of the shadow as compared to the more complicated 2D projection onto a receiver. A closed form expression for the spline shadow irradiance function is derived by twice applying Stokes’ theorem to reduce an evaluation over a 4D domain to an explicit formula involving only 2D faces on the receiver, derived from the scene geometry. This leads to a straightforward computational algorithm and an interactive implementation. Moreover, this approach can be extended to scenes involving multiple emitters and occluders, as well as curved emitters, occluders, and receivers. Spline functions are constructed from these prism-like objects. We call them generalized polyhedral splines because they extend the classical polyhedral splines to include curved boundaries and a density function. The approach can be applied to more general problems such as some of those occurring in radiosity, and other related topics.


Keywords: rendering, shadow algorithms, illumination, visibility determination

1 INTRODUCTION

While computing shadows from a single point source is relatively simple and goes back to the early days of graphics when the light source was moved away from the eye position, the accurate representation of shadows resulting from area sources is quite complicated, as Soler and Sillion [38] point out:

The calculation of detailed shadows remains one of the most difficult challenges in computer graphics, especially in the case of extended (linear or area) light sources.

In this paper we present a solution to the general differential area to polygon form factor in the presence of a polygonal occluder, by representing the induced irradiance as a spline function. We generalize the polyhedral spline [4, 16, 32] to obtain our result. The polyhedral spline is defined as a distribution. A geometric interpretation of this definition involves regarding the splines as cross-sectional volumes of higher dimensional polyhedra [14]. The more familiar univariate and tensor product B-splines are special cases of the polyhedral spline.

To introduce the fundamental ideas of our approach, we begin with an analysis in Flatland [26], where the higher dimension construction can be directly visualized and more easily understood. Then we work through the construction of a generalized perspective prism in 4D associated with an element that can be either an emitter or an occluder. The 4D perspective prism is defined so that the 2D cross-section taken through any point \((x, y)\) is a perspective view of the element from the point \((x, y, 0)\). A shadow function can be built by considering boolean operations on such prisms. In the case where the emitter \(E\), occluder \(O\), and receiver \(R\) are all parallel, the intersection region of the \(E\) and \(O\) prisms is isomorphic to the 4D Cartesian product \(O \times E\), where \(O\) and \(E\) are taken as 2D polygons. To bolster confidence for looking into 4D, we observe that the projection of \(O \times E\) onto \(R\) generates all the discontinuity lines of the shadow structure.

The boolean combination of the \(O\) and \(E\)-prisms gives a polyhedron-like object in 4D which is used to define a polyhedral spline function. At each point \((x, y)\) in the plane the value of the spline is obtained by integrating a density function over the 2D cross-section at \((x, y)\). With a proper definition of the density function this spline gives us the exact partially occluded irradiance at \((x, y)\). Using Stokes’ theorem we derive a recurrence relation that yields a closed form solution for the irradiance function.

1.1 Related Work

Accurate computation of shadows has been a rich and active area of graphics research since near the beginning of the field. In 1990 Woo et al. [44] made an extensive survey of relevant techniques. Since then many techniques have been aimed at faster or more accurate computation.

Methods that are based on a single point light source create hard shadows. A typical example of such a method is used in ray tracing, where a ray is cast between a point lying on the surface and a specified light source. The algorithm must determine if there is an occluder between it and the light source [42]. Soft shadows caused by area light sources are simulated by stochastically selecting a set of sample points on the source and casting several rays towards them from the surface [11, 36]. Exact evaluation can be obtained by clipping all the occluders against the area source [18, 33].

Several methods consider the view from a point light source. The shadow buffer creates a discrete depth image from the point light source view [43], using only a single point. Shadow volumes are constructed from a point light source and polygonal occluders [13] or curved boundary occluders [27], from which a point can be tested for its visibility in object space. Soft shadows can be approximated in a similar manner to the ray tracing method, by breaking the light
source into several point samples [7].

The radiosity method theoretically includes the computation of soft shadows. However, the method requires the reconstruction of irradiance functions, for which no general closed form is known. Methods to compute radiosity use a mesh of finite surface elements, then “form factors” representing the exchange of energy between mesh elements are computed. Frequently, the implementation assumes a constant radiosity distribution across each surface patch and finite element, then the values are interpolated across each mesh element. Meshing surfaces without considering the shadow structure can lead to inaccurate solutions and “shadow leaks”. The discontinuity meshing approach [17, 25, 30] was developed to avoid these problems. For each source, a discontinuity mesh is created on the surfaces (receivers). The mesh includes all lines where the illumination function has discontinuities, including the effects of boundaries of multiple occluders. Meshing concerns have been a principal driving force behind shadow research.

It was recognized quite early that solutions could be improved if the form factors could be computed analytically. Analytical solutions have been determined for specific geometries [28, 35]. Generally such solutions assume that each surface is flat and fully visible from the other. Baum et al [3] proposed computing the unoccluded point to polygon form factor analytically. To avoid violation of visibility assumptions, they proposed that each source patch be subdivided until all components were either fully visible or fully hidden from all elements in the environment.

Many current methods to compute form factors approximate them with an upper bound calculation, that is, they assume that an unoccluded radiosity kernel can be multiplied against a visibility factor, which gives the fraction of the area of the light source element visible from the receiver [9, 24, 40]. The point to polygon form factor is assumed to have constant exitance on the source. Recently Soler and Sillion [38] showed that when the elements are not parallel to each other, the method requires approximating the elements by parallel elements, called virtual geometry. Since the error can be estimated, it is suggested that when the error is too large, the elements should be subdivided and new orientations and approximating parallel sub-elements should be chosen until the error is within bounds. They compute irradiance using the radiosity kernel-visibility factor approximation.

Polyhedral splines have been well studied for regular polyhedra like simplices, boxes, and cones; de Boor and Höllig’s original paper [4] contains references and a recurrence relation valid for more general polyhedra than simplices, boxes, and cones. Simplex and box splines have been used in geometric modeling [19]. McCool has used polyhedral splines for signal processing and rendering analytically antialiased polygons, and also suggested they might be used for penumbral approximation [31, 32].

2 FLATLAND

We begin by carrying through our formulation in Edwin Abbott’s “Flatland.” In Flatland, surfaces become plane curves, planes become lines, and polygons become line segments [26]. Irradiance is computed from the integral

$$I(r) = \int_{C} I(r, \theta) \cos \theta \, d\theta$$

(1)

where $I(r, \theta)$ is the irradiance at the origin as a function of angle $\theta$, and $C$ is the open semi-circle above the surface at $r$. The transformation and resulting objects are all functions of $t$.

Figure 1: A general irradiance situation in Flatland. To compute the irradiance at $r(t)$, we transform to the origin and rotate so that the surface normal is $(0, 1)$. Then the transformed emitter is projected onto the line $y = 1$, where it is a line segment projectively equivalent to the apparent silhouette of $T(E)$. The transformation and resulting objects are all functions of $t$.

2.1 Formulation of Irradiance

Consider a receiver surface $R$ described by a plane parametric curve $r(t) = (x(t), y(t))$ and a normal vector $n(t) = (n_x(t), n_y(t))$. Let $E$ be a uniformly emitting surface, situated above the receiver $R$, as shown in Figure 1. In order to compute the irradiance at a point $r(t_0) = (x_0, y_0)$ on $R$ corresponding to a parameter value $t = t_0$ we proceed as follows:

1. Transform the scene such that $(x_0, y_0)$ is mapped to the origin, and such that the unit normal $n(t_0)$ becomes parallel to the $y$-axis. This transformation, which we call $T^*$, is a translation followed by a rotation, which transforms $E$ into a new position $T(E)$.

2. Let $E^*$ be the perspective projection of $T(E)$ onto the line $y = 1$ with the origin as the center of projection.

We then have

$$E^* = \{ (x, 1) : p_i^*(t) \leq x \leq p_2^*(t) \}$$

where

$$p_i^*(t) = \frac{w_{i,1}(t)}{w_{i,2}(t)}, \quad i = 1, 2,$$

and $w_i(t)$ are the silhouette points of $T(E)$ as seen from the origin. The object $T(E)$ and the projection $E^*$ produce the same irradiance at the origin, which is the irradiance at $r(t)$. Denoting this irradiance $I_{E^*}(r(t))$ by $I(t)$, we have

$$I(t) = K \int_{p_1^*(t)}^{p_2^*(t)} \frac{\cos \theta \cos \theta'}{d(x)} \, dx$$

where $K$ is a constant, $d(x)$ the distance between the origin and the point $(x, 1)$, $\theta$ the angle between the ray $(0,0) \to (x,1)$ and the $y$-axis, and $\theta'$ the angle between the same ray and the normal to $E^*$. 
Since \( E' \) is normal to the \( y \)-axis we have \( \theta' = \theta, \cos \theta = 1/d(x) \), and \( d(x)^2 = 1 + x^2 \). It follows that

\[
I(t) = K \int_{p^*_y(t)}^{p^*_x(t)} \frac{dx}{(1 + x^2)^{3/2}} = K \left[ \Omega(p^*_y(t)) - \Omega(p^*_y(t)) \right],
\]

where

\[
\Omega(x) = \frac{x}{\sqrt{1 + x^2}}.
\]

The function \( I(t) \) can be interpreted geometrically. Suppose for simplicity that \( p^*_x(t) \geq p^*_y(t) \) for all \( t \). We consider the integral

\[
\int_{p^*_y(t)}^{p^*_x(t)} \omega(y) dy,
\]

where

\[
\omega(y) = \frac{1}{(1 + y^2)^{3/2}}.
\]

Let \( M_E \) be the 2-dimensional strip

\[
M_E = \{ (t,y) : -\infty < t < \infty, p^*_y(t) \leq y \leq p^*_x(t) \}.
\]

For fixed \( t \) the integral is obtained by taking a vertical cut through \( M_E \) at \( t \) and integrating the density function from the lower boundary of \( M_E \) to the upper boundary of \( M_E \). Thus \( I(t) \) is the weighted length of this cut at \( t \).

### 2.2 Irradiance in the Presence of Occluders

Suppose that we extend the scene above by also including an occluder \( O \). In the same way as we did for \( E \) we obtain an infinite strip \( M_O \) corresponding to the occluder. Where the occluder strip overlaps the emitter strip, the covered portion of the emitter is not visible. The occluded irradiance can be computed from the integral of the same density function \( \omega(y) \) over the uncovered portion of the strip. This portion of the strip is the set difference \( M_E - M_O \).

If the set difference \( M_E - M_O \) is endowed with the density function \( \omega(y) \), the integral over each slice of the strip produces the unoccluded irradiance at that point, as in Figure 4(b). This approach generalizes. Suppose we have \( n \) occluders \( O_1, \ldots, O_n \) situated below \( E \). The irradiance at \( r(t) \) can be found by integrating the same density function over the set

\[
M_E - M_{O_1} - \cdots - M_{O_n}.
\]

If there are also multiple emitters, the total irradiance is computed by summing the contributions from the individual emitters; however, all other emitters must be treated as potential occluders, and there is also the issue of front-and-back culling.

The advantage of this construction is that all the local visibility calculations can be replaced by a single set operation that handles the occlusion geometry for the entire floor. Set differences may not be easy to compute in general, but in the restricted case where the receiver is a line and the objects line segments, the geometry turns out to be very clean.

### 2.3 Polygonal Objects

Suppose now \( E \) and \( O \) are one-dimensional polygons (line segments) and the receiver is the line \( y = 0 \) in Flatland. In this case, the parameterization of \( r(t) \) is simply \( r(t) = (t,0) \) and we simply use \( x \) for the position on the receiver. The transformation \( T \) becomes merely a translation by \( x \) and the subsequent projection \( E \rightarrow E' \) can be computed directly: if \( E \) has vertices \( (x'_i,y'_i) \), \( i = 1,2 \)

\[
p^*_i = \frac{x'_i - x}{y'_i}, \quad i = 1,2.
\]

Equation (2) can now be evaluated in closed form. Furthermore, each \( p^*_i \) as a function of \( x \) traces out a line, so the strip for \( M_E \) is an (unbounded) polygon.

To take occlusion by \( O \) into account, the strip \( M_O \) is constructed for \( O \). The partially-occluded irradiance can be computed from

\[
I(x) = S_{E'}(x) - S_{E \cap O'}(x).
\]
where

\[ S_Q(x) = \int_{y(x)}^{y(1)} \omega(y) dy \]

in case \( Q \) is the line segment from \( (a(x), 1) \) to \( (b(x), 1) \).

The value of \( S_{E_\cap M} (x) \) is therefore the integral of the irradiance density function \( \omega(y) \) over the cut at \( x \) of the intersection of the strips \( M_O \cap M_E \). This intersection has four vertices, one for each pair of edges of \( M_O \) and \( M_E \). In the special case where \( O \) and \( E \) both happen to be parallel to the \( x \)-axis, \( M_O \) and \( M_E \) are both strips of constant width and the intersection \( M_O \cap M_E \) is a parallelogram as shown in Figure 2. In the more general case, the intersection \( M_O \cap M_E \) can be shown to be a quadrilateral, as in Figure 5. The quadrilateral is not self-intersecting as long as the line containing \( E \) do not intersect \( O \), and vice-versa.

The formulation of irradiance as the integral over a cut of a two-dimensional object casts it as a weighted polyhedral spline. The constructions in this subsection show that both unoccluded and partially occluded irradiance have this formulation. What is the advantage? First of all, the formulation of irradiance in terms of a polyhedron encodes all the visibility information from the receiver in to a single, easily described two-dimensional object, as suggested in Figure 6. Second, and more importantly, polyhedral splines can be evaluated in closed form without actually finding the intersection of the cut [4].

3 THE PERSPECTIVE PRISM

In this section we construct the perspective prism analogous to the one in flatland, which stores the apparent geometry of an arbitrary polygon in \( \mathbb{R}^2 \) at every point on the floor \( z = 0 \). The construction pertains to a scene element \( Q \) which is either an emitter or an occluder.

3.1 Definition by Cross-Section

Suppose \( Q \) is a planar polygon having vertices \( v_i = (x_i, y_i, z_i) \) for \( i = 1, \ldots, n \). Furthermore, assume \( Q \) is situated strictly above the floor, so that each \( z_i > 0 \). We define a subset \( M_Q \) of \( \mathbb{R}^4 \) in terms of its 2D cross-sections \( P_{(x,y)} \cap M_Q \), where for fixed \( (a,b) \)

\[ P_{(a,b)} = \{(a,b,z,w) : z, w \in \mathbb{R}\} \]

is an \((a,b)\)-slice.

For each \((x,y)\) the cross-section is given by the polygon \( P_{(a,b)} \cap M_{Q} \) including its interior, where

\[ V_i = (x, y, \frac{x - x_i}{z_i}, \frac{y - y_i}{z_i}), \quad i = 1, \ldots, n. \]  

We observe that the last two components of \( V_i \) are the perspective transformation of \((x, y)\) with the origin moved to \((x, y)\).

For later use we define \( Q^* \) as the polygon with vertices \(((x_i - x)/z_i, (y_i - y)/z_i, 1)\) for \( i = 1, \ldots, n \). Thus \( Q^* = Q^*(x, y) \) is a perspective transformation of \( Q \) including its interior to \( z = 1 \) with the origin moved to \((x, y)\).

3.2 The Geometric Structure of the Perspective Prism

Describing the cross-sections of \( M_Q \) is sufficient to properly define the prism as a point set, but the slices give only a local characterization, and, for our purpose, we need a quantitative understanding of the boundary of all of \( M_Q \).

Suppose \( i \) is fixed. As \((x, y)\) varies, the \( i \)th vertex \( V_i = V_i(x, y) \) can be interpreted as a 2-surface in \( \mathbb{R}^4 \) that is in fact a 2-plane.

Each apparent edge of \( Q \) is a line segment that is a function of \((x, y)\), and can be parameterized by

\[ E_i(x, y, t) = (1 - t)V_i(x, y) + tV_{i+1}(x, y). \]
This shows that therefore the union of the sweep-surfaces of all the apparent edges. The apparent emitting polygon is simply with the domain of $M$. In this case, $z = h$ for each $i$, where $h$ is the height of $Q$ above the floor, and the vertex surface functions $V_i$ becomes

$$\begin{bmatrix} x \\ x_i - x \\ y_i - y \\ z_i \\ \frac{y_i - y}{z_i} \end{bmatrix} = \frac{1}{h} \begin{bmatrix} 0 & 0 & x_i & 0 & \frac{x}{h} \\ 0 & 0 & -x_i & 0 & -\frac{x}{h} \\ y_i & 0 & 0 & \frac{y}{h} & 0 \\ -y_i & 0 & 0 & -\frac{y}{h} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ y_i \\ x_i \\ -x_i \end{bmatrix}$$

Thus $M_Q$ can be viewed as $Q$ rotated so that it is parallel to the $zw$-plane, scaled by $1/h$, and swept along a 2-plane in $\mathbb{R}^4$. The edge surface functions, from (4)

$$E_i(x, y, t) = \begin{bmatrix} x \\ x_i - x \\ \frac{y_i - y}{h} + \frac{y_i - y}{h} \\ z_i \\ \frac{z_i - z}{h} + \frac{z_i - z}{h} \end{bmatrix}$$

with the domain of $t$ extended to all of $\mathbb{R}$, become proper hyperplanes (affine images of $\mathbb{R}^3$) in $\mathbb{R}^4$. Therefore the perspective prism $M_Q$ is a genuine, albeit unbounded, four-dimensional polyhedron.

## 4 THE OCCLUDER AND EMITTER PRISM

We now proceed to the situation where there is an occluding polygon in the scene. Suppose we have in $\mathbb{R}^3$ a uniformly emitting polygon $E = \{e_1, \ldots, e_m\}$ and an occluding polygon $O = \{o_1, \ldots, o_n\}$. We assume that $E$ lies strictly above the plane of $O$, and also that both $E$ and $O$ lie strictly above the floor.

### 4.1 Algebra of Perspective Prisms

To analyze partially occluded emitters we need a 4D object $M$ whose $(x, y, t)$ cross-sections store the geometry of the visible portion of the emitter. At each floor point $(x, y)$ the visible portion of the apparent emitting polygon is simply $E^* - O^*$, where $E^*$ and $O^*$ are $E$ and $O$ projectively transformed through $(x, y, t)$. From the definition of the perspective prism it follows immediately that

$$M = M_E - M_O.$$  

Dealing with occlusion is now reduced to a tractable problem of performing booleans of the prisms.

### 4.2 The Structure of $O \times E$

If $E$ and $O$ are both parallel to the floor, then $M_i$ and $M_E$ are both four-dimensional “prism” polyhedra, and consequently their intersection is another polyhedron. It can be shown that the intersection of the two “prisms” is isomorphic to the product polyhedron $O \times E$, as in Flatland. In the more general case of non-parallel polygons the boundary structure is topologically equivalent to $O \times E$.

The product polyhedron $O \times E$ can be constructed by sweeping $E$, rotated to the $zw$-plane across all of $O$, or vice-versa. The boundary of $O \times E$ can be viewed as $E$ swept along each edge of $O$, and vice-versa, as suggested in Figure 7. (Of course, the object is really four-dimensional, so these figures are only a representation.) Sweeping $E$ along an edge of $O$ produces a prism in the shape of $E$, a three-dimensional polyhedron embedded in $\mathbb{R}^4$. Similarly, sweeping $O$ along an edge of $E$ produces an “occluder” prism. The boundary of $O \times E$ therefore consists of $n$ emitter prisms, and $m$ occluder prisms, each a three-dimensional polyhedron situated in $\mathbb{R}^4$. These prisms are the $n \times m$ facets of $O \times E$.

Each emitter prism has two emitter faces, which are two-dimensional polygons situated in $\mathbb{R}^3$. Adjacent emitter prisms share a facet, and there is thus exactly one emitter face for each occluder prism, and vice-versa. The rectangular sides of the prisms are also proper two-dimensional polygons situated in $\mathbb{R}^4$. Occluder prism $j$ shares exactly one side with emitter prism $i$, hence there are $m \times n$ sides. The $n$ emitter faces, the $m$ occluder faces, and the $m \times n$ sides together form the boundaries of the facets of $O \times E$. 

---

*Figure 7: The product polyhedron $O \times E$ is formed by sweeping a copy of $E$ (the red triangle) across $O$ (the blue pentagon) or vice-versa. The boundary is thus formed by sweeping each object along each edge of the other, resulting in three-dimensional prisms. Occluder prism $j$ shares a rectangular “side” with emitter prism $i$.***

*Figure 8: The five emitter-prism facets of $O \times E$. Adjacent emitter prisms share a common emitter face.*

*Figure 9: The three occluder-prism facets of $O \times E$. Adjacent occluder prisms share a common occluder face.*
5 IRRADIANCE

Our goal is to build a shadow function that produces exact irradiance everywhere. We review what the irradiance quantity is for a scene with polygonal occluders and emitters.

We recall that the irradiance from a uniformly emitting surface \( S \), which is not self-occluding as viewed from a point \((x_0, y_0)\) on a receiver, can be computed [1] from the surface integral

\[
I(x_0, y_0) = K \int_S \frac{\cos \theta \cos \phi}{|r(x, y)|^2} dS,
\]

where \( r(x, y) = (x-x_0, y-y_0) \) is the ray from \((x_0, y_0)\) to a point \((x, y)\) on \( S \), \( \theta \) is the angle between the normal to the receiver at \((x_0, y_0)\) and \( r(x, y) \), while \( \phi \) is the angle between \( r(x, y) \) and the normal to \( S \) at \((x, y)\). The constant \( K \) is a physical parameter of \( S \). For simplicity we hereafter assume that \( K = 1 \).

When the emitting surface is a polygon, the integral in (6) can be shown to be equal to

\[
I(x_0, y_0) = \int_{E^*} \frac{1}{(1 + x^2 + y^2)^2} dx \, dy
\]

(7)

where \( E^* = E^*(x_0, y_0) \) is the perspective transformation of \( E \) onto the plane \( z = 1 \) with the origin moved to \((x_0, y_0)\). For each polygon \( Q^* \) in the plane \( z = 1 \) we define the function

\[
S_{Q^*} = \int_{Q^*} \frac{1}{(1 + x^2 + y^2)^2} dx \, dy.
\]

(8)

In an unoccluded scene with a emitter \( E \) the irradiance at a point \((x_0, y_0)\) on the receiver is simply \( I(x_0, y_0) = S_{E^*}(x_0, y_0) \). In a scene in which additional has an occcluder \( O \) it follows from the algebra of perspective prisms that

\[
I(x_0, y_0) = S_{E^* \cap O^*} = S_{E^*} - S_{E^* \cap O^*}.
\]

(9)

This is another way of saying that the net irradiance at the point \((x_0, y_0)\) on the receiver is obtained by subtracting the occluded irradiance from the full unoccluded quantity.

6 GENERALIZED POLYHEDRAL SPLINES

In this section we develop some mathematical tools that are needed to describe shadows. Suppose \( B \) is a bounded convex polyhedron body in \( \mathbb{R}^n \) whose boundary \( \partial B \) is a disjoint union of finitely many hyperplanes \( F_i \), and suppose \( \omega = \omega(y) = \omega(y_1, \ldots, y_n) \) is a function of \( n \) variables which is continuous and integrable over \( B \). Given an integer \( s \leq n \) let \( Q \) be the projection \( \mathbb{R}^n \to \mathbb{R}^s \) given by \( P(y_1, \ldots, y_n) = (y_1, \ldots, y_s) \). We define a generalized polyhedral spline \( N_{B,\omega}(x) \) of \( s \) variables such that for all sufficiently smooth functions \( \phi \) of \( s \) variables with local support

\[
\int_{\mathbb{R}^s} N_{B,\omega}(x) \phi(x) \, dx = \int_{B} \omega(y) \phi(P(y)) \, dy.
\]

(10)

Note that (10) defines \( N_{B,\omega} \) as a distribution which we assume can be defined as a continuous function. If \( n = s \) we then have

\[
N_{B,\omega}(x) = \chi_B(x) \omega(x),
\]

(11)

where \( \chi_B(x) = 1 \) if \( x \in B \) and \( 0 \) otherwise.

If \( \omega = 1 \) the function \( N_{B,\omega} \) is the multivariate polyhedral spline [4, 32]. We will use more general densities \( \omega \) for the definition of our shadow irradiance functions.

By writing the right-hand side of (10) as an iterated integral we obtain a common geometric interpretation of \( N_{B,\omega} \)

\[
N_{B,\omega}(x) = \int_{\mathbb{R}^s} \chi_B(x, z) \omega(x, z) \, dz.
\]

(12)

In this interpretation the value of the spline at a point \( x \in \mathbb{R}^s \) is obtained by integrating the density function \( \omega \) over a \( n - s \)-dimensional cross-section of \( B \) located at \( x \). This cross-section is the intersection of \( B \) and an \( x \)-slice in \( \mathbb{R}^n \).

To make the above formula tractable for evaluation we apply Stokes’ theorem to (10) and obtain

\[
N_{B,\omega}(x) = \sum_i \text{sgn} \left( v_i \right) N_{B,\phi_i}^{s-1}(f_i) \omega(\Phi_i(x)).
\]

(13)

Here \( F_i \) is facet \( i \) of \( B \), having outward normal \( n_i \), \( v_i \) is the \( n \)-th component of \( n_i \), \( \text{sgn}(t) \) is equal to \( 1, 0, -1 \) if \( t > 0, t = 0, t < 0 \), respectively, \( \Omega \) is an anti-derivative of \( \omega \) with respect to the last variable, \( P_{s-1}^r(F_i) \) is the canonical projection of \( F_i \) to \( \mathbb{R}^{s-1} \), and \( \Phi_i \) is a function of \( n - 1 \) variables describing \( F_i \) in the form \( \Phi_i(y_1, \ldots, y_{s-1}) = (y_1, \ldots, y_{s-1}, y_s) \), where \( y_s \) is an equation for the hyperplane \( F_i \) in term of \( y_1, \ldots, y_{s-1} \). The \( \text{sgn} \) function handles a reversal of the orientation of facet \( F_i \) in the projection to the lower space.

We note that the functions on the right of (13) are splines of \( s \) variables defined using objects in \( \mathbb{R}^{s-1} \). We can apply this formula recursively to reduce \( n \). The recurrence stops when \( n \) reaches \( s \), in which case the spline is given by (11).

The unwighted polyhedral spline recurrence (\( \omega = 1 \)) developed by de Boor and Hollig [4], as well as the polynomially weighted recurrence of McCool [32] are both derived following an identity of Hakopian. Ours takes a different approach, based directly on Stokes’ theorem. The main use for our formula is to derive the explicit shadow irradiance function in the next section.

6.1 Extension to Non-Polyhedral Objects

The recurrence (13) can be extended beyond proper polyhedra, to certain general manifolds having the same general structure as polyhedra, inasmuch as their boundaries consist of facet-like manifolds, each of one lower dimension. As we shall see, the general occluded polygonal irradiance problem leads to manifolds of this form. The primary difficulty of this extension is that the “facets” may not be expressible as explicit surfaces of their projections, and when they are, they are no longer the simple affine hyperplane functions as in the polyhedron case. Problems such as this can sometimes be circumvented by partitioning the boundary facets into disjoint pieces; however, in the object we are about to develop, it happens that we can effectively ignore this potential problem.

6.2 The Irradiance Function

In Section 5 we gave an explicit formula for the irradiance functions based on the geometry of prisms in \( \mathbb{R}^s \). In this section we have developed the appropriate spline theory to construct the irradiance functions. The remaining task is to combine the two approaches to show that the irradiance function is, in fact, a spline. First consider a scene with an unoccluded emitter. The irradiance at a point \((x_0, y_0)\) on the receiver is given in (7), which we write in the form

\[
I(x_0, y_0) = \int_{E^*} \omega(x_0, y_0, z, w) \, dw \, dz,
\]

with

\[
\omega(x_0, y_0, z, w) = \frac{1}{(1 + z^2 + w^2)^2}.
\]

(14)
Figure 10: The shadow discontinuity lines consist of $E$ projected through each point of $O$, and vice-versa.

Figure 11: The floor shadow structure for a non-parallel configuration. Here $n = 4$ and $m = 3$. (a) The three occluder images; each has a counter-clockwise orientation. (b) The four emitter images; each has a counter-clockwise orientation. (c) Two of the side images: $S_{1,1}$ has a negative (clockwise) orientation, $S_{2,1}$ is positively oriented.

Since $E'$ is essentially the cross-section of the prism $M_E$ at $(x_0, y_0)$ we obtain

$$I(x_0, y_0) = \int_{\mathbb{R}^3} \chi_{M_E}(x_0, y_0, z, w) \omega(x_0, y_0, z, w) \, dz \, dw.$$  

Hence from (12) with $n = 4, s = 2$ we obtain

$$I(x_0, y_0) = N_{M_E, \omega}(x_0, y_0),$$  

and we have demonstrated that the irradiance function is a spline. This agrees with the fact that a shadow is a piecewise smooth function with jumps in derivatives across the discontinuity lines, and the nature of spline functions. The above derivation holds equally well in the more general situation including multiple elements. For example in the situation described in (9) we have

$$I(x_0, y_0) = N_{M_{E-M_O}, \omega}(x_0, y_0).$$  

7 THE SHADOW EQUATION

Using the recurrence relation (13) we can actually give a closed-form expression for the irradiance function corresponding to a scene with a single planar polygonal emitter and a planar polygonal occluder, in terms of the space coordinates of the vertices

$$O = \{(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)\},$$

$$E = \{(x'_1, y'_1, z'_1), \ldots, (x'_m, y'_m, z'_m)\}.$$  

We assume $O$ and $E$ are convex, and $O$ lies strictly above the floor, and the plane of $O$ lies strictly below $E$. We define the direction of the normal $n_E$ of $E$ to be opposite the direction of emittance. (This is the opposite of the usual convention, but we need the emitter projections onto the plane to have counter-clockwise orientation.) The region on the plane in which the emitting side of $E$ is visible is the set

$$R_E = \{ (x, y) \in \mathbb{R}^2 : (x, y, 0) \cdot (n_E - (x_1, y_1, z_1)) > 0 \}.$$  

The formula presented in this section is valid on the planar region $R_O \cap R_E$.

The shadow equation is expressed in terms of the characteristic functions of the projected 2D boundary faces of $O \times E$. To define these polygons, first define the shadow nodes

$$\psi_{ij} = \left( \frac{x'_j - x_i}{x'_j - x_i}, \frac{y'_j - y_i}{y'_j - y_i}, \frac{z'_j - z_i}{z'_j - z_i} \right).$$  

Geometrically, $\psi_{ij}$ is the point on the floor where $o_i$ and $e_j$ appear to coincide. The shadow polygons consist of $m$ occluder images $O_j$, $n$ emitter images $E_i$, and $mn$ quadrilaterals $S_{ij}$, as shown in Figures 10 and 11. In terms of the shadow nodes,

$$O_j = \{(v_{i,1}, \ldots, v_{im})\},$$

$$E_i = \{(v_{i+1,1}, \ldots, v_{i+1,m})\},$$

$$S_{ij} = \{(v_{i+1,j}, v_{i+1,j+1}, v_{i,j+1}\}.$$  

The $z$ and $w$ coordinates of the vertices of the $(x, y)$-slices of the objects give the apparent geometry from $(x, y)$. Define (cf. Section 3.1)

$$p_i(x, y) = (p_{i,1}, p_{i,2}) = \left( \frac{x - x_i}{z_i}, \frac{y - y_i}{z_i} \right).$$

$$p'_j(x, y) = (p'_{j,1}, p'_{j,2}) = \left( \frac{x' - x'_j}{z'_j}, \frac{y' - y'_j}{z'_j} \right).$$  

It is convenient to introduce

$$e_i(x, y) = (e_{i,1}, e_{i,2}) = p_{i+1}(x, y) - p_i(x, y),$$

$$e'_j(x, y) = (e'_{j,1}, e'_{j,2}) = p'_{j+1}(x, y) - p'_j(x, y).$$  

For the evaluation of the splines, the function defining the $i$th facet must be written in the form $w_i(x, y, z) = m_i(x, y) z + b_i(x, y)$, where for $O$,

$$m_i(x, y) = \frac{e_{i,2}}{e_{i,1}},$$

$$b_i(x, y) = p_{i,2} - \frac{e_{i,2}}{e_{i,1}} p_{i,1},$$  

and for $E$,

$$m'_j(x, y) = \frac{e'_{j,2}}{e'_{j,1}},$$

$$b'_j(x, y) = p'_{j,2} - \frac{e'_{j,2}}{e'_{j,1}} p'_{j,1}.$$  

The second iteration of (13) requires the equations for the boundary of the facets of $M_O$, which are the 2-surfaces traced out by the vertices of $O^*$ and $E^*$ in $M_O$ and $M_E$. These are essentially given by (3). There are also the extra boundary 2-surfaces traced by the intersection of edges $i$ of $O^*$ and $j$ of $E^*$, which can be computed from

$$\psi_{ij}(x, y) = \frac{e_{i,1} e'_{j,2} p_{j,1} - e'_{j,1} (e_{i,2} p_{j,1} + e_{i,1} (p_{j,2} - p_{i,2}))}{e_{i,1} e'_{j,2} - e_{i,2} e'_{j,1}}.$$  

Finally, we need the twice-integrated density function

$$F(z, u, v) = \frac{u \arctan \left( \frac{u + \sqrt{1 + u^2}}{v + \sqrt{1 + v^2}} \right)}{2 \sqrt{1 + u^2 + v^2}} + \frac{z \arctan \left( \frac{u + \sqrt{1 + u^2}}{v + \sqrt{1 + v^2}} \right)}{2 \sqrt{1 + u^2 + v^2}}.$$  

The irradiance on the plane $z = 0$ is then

$$I(x, y) = N_{M_E, \omega}(x, y) - N_{M_O \cap M_E, \omega}(x, y).$$
where \( \omega \) is given by (14). Applying (13) twice it can be shown that

\[
N_{M_j \cap M_j \cap \omega}(x, y) = \sum_{j=1}^{m} \chi_{\tilde{O}_j}(x, y) \left[ F(p_{j,1}, m_j, b_{j}) - F(p_{j,1}, m_j, b_{j-1}) \right] \\
+ \sum_{i=1}^{n} \chi_{E_i}(x, y) \left[ F(p_{i,1}, m_i, b_i) - F(p_{i,1}, m_{i-1}, b_{i-1}) \right] \\
+ \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha(S_{ij}) \chi_{S_{ij}}(x, y) \left[ F(\psi_{ij}, m_i, b_i) - F(\psi_{ij}, m_j, b_j) \right].
\]

(23)

Here \( \tilde{Q} \) is the polygon \( Q \) excluding edges whose oriented angle with the \( x \)-axis is greater than zero, and less than \( \pi \). The \( \alpha(Q) \) function is the orientation of the polygon \( Q \) : \( \alpha(Q) \) is 1 if the vertices are in counter-clockwise order, and -1 otherwise.

Inspection of the equation for \( N_{M_j \cap M_j \cap \omega}(x, y) \) reveals the first summation is in fact the unoccluded irradiance, hence the irradiance can be expressed directly in the following shadow equation.

\[
I(x, y) = \sum_{j=1}^{m} \chi_{\tilde{O}_j}(x, y) \left[ F(p_{j,1}, m_j, b_{j}) - F(p_{j,1}, m_j, b_{j-1}) \right] \\
- \sum_{i=1}^{n} \chi_{E_i}(x, y) \left[ F(p_{i,1}, m_i, b_i) - F(p_{i,1}, m_{i-1}, b_{i-1}) \right] \\
- \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha(S_{ij}) \chi_{S_{ij}}(x, y) \left[ F(\psi_{ij}, m_i, b_i) - F(\psi_{ij}, m_j, b_j) \right].
\]

(24)

The shadow spline function \( I(x, y) \) is therefore a collection of smooth functions supported on the shadow polygons \( O_j \), \( E_i \) and \( S_{ij} \).

Images demonstrating the formula are shown in Figures 12, 13.

These images were rendered using ray tracing, with the irradiance on the surfaces calculated directly using the shadow equation.

The formula extends to multiple occluders as long as their shadows do not interact. The facets of a convex polyhedral occluder meet this, and the requirements of the shadow equation (if the normals point inward), so the formula can be extended to convex polyhedral emitters and occluders, as demonstrated in Figures 14 and 15.

8 FUTURE WORK AND CONCLUSION

The general formulation presented in this paper is rich in possibilities worthy of deeper investigation. As a first step, we hope to generalize the non-interfering occluder formula to take into account multiple objects and their interactions with each other. Beyond this, we are encouraged by the possibility of producing an analytic formula for objects more general than polygons, and receivers more general than portions of a plane. Furthermore, the ideas generalize to non-uniform emittance and arbitrary BRDFs.

8.1 Conclusion

In this paper we analyze the essential mathematical structure of shadows on a surface and show that the natural place to understand them is in a higher dimensional space. We show that the general problem of calculating the irradiance function for the shadow cast by a single polygonal occluder \( O \) from a polygonal emitter \( E \) in general position leads to a weighted (generalized) polyhedral spline function, and that the Cartesian product \( O \times E \) provides the structure necessary to compute the shadow. To compute the irradiance for non-parallel configurations of \( E \) and \( O \), ordinary polyhedral splines must be generalized. The resulting splines are built on objects which are deformations of polyhedra in higher dimensions, instead of the usual flat-faced structures.
Using Stokes’ theorem to reduce the Euclidean dimension of the integration problem, we arrive at a closed-form formula for evaluating the irradiance function, not just the partial visibility quantity, as an exact analytical expression. The wire frame structure of the 4D object by itself produces the shadow discontinuity lines when projected to the floor. That is, the projection method naturally specializes to the produce the discontinuity graph described by Nishita and Nakamae [33]. When $E$, $O$, and the floor are all parallel and a constant weighting function is specified, the spline specializes to an $O \times E$ (prism) polyhedral spline that yields the convolution presented by Soler and Sillion in their shadow algorithm [38]. Although it is a common simplifying approximation to factor out visibility from the irradiance integral, there is no penalty in our method for doing it exactly.

Although exact shadows may not be necessary in many situations, they are important because their structure gives fundamental understanding. Knowing their exact behavior also provides a way to compute the error in approximate methods. And finally, the direct evaluation of the explicit formula is fast enough to compete with approximate algorithms.

A benefit of studying shadows in the higher dimensional context with the constructions provided is that the approach generalizes beyond the solution for a single emitter and occluder in arbitrary configuration. This approach also generalizes theoretically in a straightforward manner to more complicated situations involving curved and multiple emitter and occluder elements, at the expense of complicating the geometry. To carry forward this approach for curved and nonuniform elements we must again use generalized polyhedral splines.

Computing irradiance quantities for shadow functions is equivalent to computing the differential area to polygon form factor in the presence of an occluder. Although we do not take the space to develop the details, one can see that this approach may be used in radiosity. There is also a connection between the 4D emitter or occluder prisms and some image based rendering algorithms.

Considering shadow irradiance functions as higher dimension generalized polyhedral splines has lead to several advances in understanding and computation. Although somewhat abstract, the approach described herein, which appears to be quite general and powerful, seems rather promising as a basis from which to do further work in this area and related problems.

Acknowledgments

This work was supported in part by DARPA (F33615-96-C-5621) and the NSF Science and Technology Center for Computer Graphics and Scientific Visualization (ASC-89-20219). The authors wish to thank John Hughes, Brian Smits, and Peter Shirley for reading earlier drafts of this paper, and the reviewers for their comments.

References
