A Continuous Adjoint Formulation for Radiance Transport

Per H. Christensen David H. Salesin Tony D. DeRose

Department of Computer Science and Engineering
University of Washington
Seattle, Washington 98195

March 26, 1993

1 Introduction

Radiosity algorithms assume that all reflection is ideally diffuse. This assumption, while making the computation of global illumination more tractable, ignores many important effects, such as mirror reflection and glossy highlights. Though more expensive, the simulation of directional distributions of lights, known as *radiance transport*, is essential for realistic image synthesis.

To date, two general approaches have been used for radiance transport: Monte Carlo and finite-element methods.

The Monte Carlo approach, as pioneered by Cook *et al.* [1], Kajiya [4], and Ward *et al.* [9], samples the directional distributions using "random walks," which in effect simulate the transport of light by following paths of sample photons. Error is easy to control: as more rays are traced, the variance of the solution is reduced. However, Monte Carlo approaches, while easy to implement and control, are typically slow to converge—and are particularly ill-suited to diffuse reflectors, which require a large number of rays to sample adequately.

The finite element approach, as pioneered by Immel *et al.* [3], Shao *et al.* [6], and Sillion *et al.* [7], discretizes each surface into small elements and approximates the radiance over each element by some function. This discretization can be thought of as a "projection" of the original continuous (infinite-dimensional) problem into a finite-dimensional domain. The radiance solution in this projected space is then given by a linear system of equations. The approach is *view-independent* in the sense that the radiance solution, once computed, is equally valid for any eyepoint and view direction.

For n elements, the finite-element approach gives rise to a system of $O(n)$ equations with $O(n)$ unknowns, which in general requires $O(n^2)$ storage and $O(Kn^2)$ time to solve, where $K \leq n$ is a constant that depends on the reflectivity of surfaces in the scene. Since for graphics applications n can be in the thousands or even millions, these solution requirements can be prohibitive. Recently, Hanrahan *et al.* [2] proposed a hierarchical algorithm for the simpler case of radiosity that goes a long way toward resolving this difficulty. The hierarchical algorithm focuses effort on the significant energy transfers, quickly approximating the insignificant interactions.

An even larger speedup can be obtained by observing that a view-independent solution is not always necessary; in many cases, we would be satisfied with an accurate solution for a particular view or set of views. To address this issue, Smits *et al.* [8] described a radiosity algorithm that extends the hierarchical approach to refine just those interactions contributing the most error to a *view-dependent* solution. The algorithm makes use of *importance functions*, which are defined as the solution to the adjoint radiosity transport equation. The "importance" of a given patch essentially describes how radiosity originating at that patch influences the visible surfaces. The algorithm combines estimates of importance and radiosity to drive the global solution, allowing it to exploit view-dependent information as part of an adaptive refinement scheme.

In this paper, we show how the adjoint formulation for radiosity can be extended to the more general setting of radiance transport, where the payoff from using importance is even greater. This speedup is due to the savings in computation for highlights that are present in the visible scene, but not directly visible to the eye. Thus, in contrast to radiosity transport, applying importance to radiance transport can give dramatic speedups even when all elements of the scene are visible. Seen another way, "radiosity space" is two-dimensional (one radiosity for every surface point), whereas "radiance space" is four-dimensional (one radiance for every surface point and direction), so we might reasonably expect that limiting our error metric to a small subset of this higher-dimensional space will yield a greater payoff.

The contributions of this paper are several. We give a continuous adjoint formulation of radiance transport, which extends the discrete adjoint formulations of radiosity transport by Smits *et al.* [8] and of radiance transport by Pattanaik and Mudur [5]. We observe that the angular distributions of radiance are in general continuous functions, whereas the angular distributions of its adjoint, "directional importance," are in general discontinuous. This observation motivates the formulation of a related, continuous type of directional importance, which we prove to be equivalent to radiance in the sense that the two quantities satisfy the same transport equation. This form of directional importance can be propagated through the environment in exactly the same fashion as radiance. Finally, we present results from a preliminary implementation that demonstrate the potential of using an adjoint formulation to speed the computation of a view-dependent radiance transport solution.

2 Radiance

Let x, y, and z be points in space. We define the *radiance* $\mathrm{L}(y \to z)$ as the power emanating from y, per unit solid angle in the direction $z - y$, per unit projected area perpendicular to $z - y$. Radiance L is measured in [watt$\cdot$meter$^{-2}\cdot$steradian^{-1}]. The transport of radiance is described by the following equation:

$$\mathrm{L}(y \to z) \;= \overset{\circ}{\mathrm{L}}(y \to z) + \int_x f_r(x \leftrightarrow y \leftrightarrow z)\, G(x \leftrightarrow y)\, \mathrm{L}(x \to y)\, dx. \qquad (1)$$

In this equation, $\overset{\circ}{\mathrm{L}}(y \to z)$ is the *emitted radiance* from y in direction $z - y$. It has the same units as radiance [watt $\cdot$ meter^{-2} $\cdot$ steradian^{-1}].

The term $f_r(x \leftrightarrow y \leftrightarrow z)$ is the *bidirectional reflectance-distribution function*, or BRDF, and describes the ratio of reflected radiance in direction $z - y$ to the differential irradiance from direction $y - x$ that produces it. As a consequence of Helmholtz reciprocity, the BRDF satisfies $f_r(x \leftrightarrow y \leftrightarrow z) = f_r(z \leftrightarrow y \leftrightarrow x)$; we therefore use double-arrows ($\leftrightarrow$) between its arguments. The BRDF has units [steradian^{-1}].

Finally, the *geometric term* $G(x \leftrightarrow y)$ is given by

$$G(x \leftrightarrow y) \;\equiv\; V(x \leftrightarrow y) \cdot \frac{\cos\theta_x \cos\theta_y}{||x - y||^2},$$

where $V(x \leftrightarrow y)$ is a *visibility term* that is 1 or 0, depending on whether or not x and y are visible to one another, and θ_x and θ_y are the angles between line segment xy and the respective normals to differential areas at x and y. The geometric term describes how radiance leaving a differential area at x in direction $y - x$ arrives as differential irradiance at y. It has units [steradian $\cdot$ meter^{-2}].

The radiance transport equation can be rewritten as $\mathrm{L} = \overset{\circ}{\mathrm{L}} + \mathcal{T}\mathrm{L}$ or

$$(\mathcal{I} - \mathcal{T})\mathrm{L} = \overset{\circ}{\mathrm{L}}, \tag{2}$$

where $\mathcal{I}$ is the identity operator, and the *transport operator* $\mathcal{T}$ is defined as

$$(\mathcal{T}\mathrm{L})(y \to z) \equiv \int_x f_r(x \leftrightarrow y \leftrightarrow z)\, G(x \leftrightarrow y)\, \mathrm{L}(x \to y)\, dx. \tag{3}$$

3 Directional importance

Two operators $\mathcal{O}$ and $\mathcal{O}^*$ are said to be *adjoint* if and only if for all A and B,

$$\langle \mathcal{O}A\,,\, B\rangle = \langle A\,,\, \mathcal{O}^* B\rangle \tag{4}$$

for some inner product $\langle \cdot\,,\, \cdot \rangle$.

Let $\Gamma(x \leftarrow y)$ be a new quantity, called *directional importance*, that is transported by the adjoint equation of (2),

$$(\mathcal{I} - \mathcal{T})^*\Gamma = \overset{\circ}{\Gamma}, \tag{5}$$

where the adjoint is defined with respect to the inner product

$$\langle A\,,\, B\rangle \equiv \int_{xy} A(x \to y)\, B(x \leftarrow y)\, dy\, dx.$$

Lemma *The adjoint transport operator $\mathcal{T}^*$ is*

$$(\mathcal{T}^*\Gamma)(x \leftarrow y) = G(x \leftrightarrow y) \int_z f_r(x \leftrightarrow y \leftrightarrow z)\, \Gamma(y \leftarrow z)\, dz. \tag{6}$$

Proof We need to verify that $\mathcal{T}$ and $\mathcal{T}^*$ are adjoints in the sense of equation (4). This follows from simple manipulation of integrals:

$$\langle \mathcal{T}\mathrm{L}\,,\, \Gamma\rangle = \int_{zy} (\mathcal{T}\mathrm{L})(y \to z)\, \Gamma(y \leftarrow z)\, dy\, dz$$

$$
\begin{aligned}
&= \int_{zy}\int_{x} f_r(x \leftrightarrow y \leftrightarrow z)\, G(x \leftrightarrow y)\, \mathrm{L}(x \to y)\, dx\ \Gamma(y \leftarrow z)\, dy\, dz \\
&= \int_{xy} \mathrm{L}(x \to y)\, G(x \leftrightarrow y) \int_{z} f_r(x \leftrightarrow y \leftrightarrow z)\, \Gamma(y \leftarrow z)\, dz\, dy\, dx \\
&= \int_{xy} \mathrm{L}(x \to y)\, (\mathcal{T}^*\Gamma)(x \leftarrow y)\, dy\, dx \\
&= \langle \mathrm{L}\,,\, \mathcal{T}^*\Gamma \rangle
\end{aligned}
$$

□

Theorem *The adjoint equation for radiance transport* (1) *is*

$$\Gamma(x \leftarrow y) \;=\; \mathring{\Gamma}(x \leftarrow y) + G(x \leftrightarrow y) \int_{z} f_r(x \leftrightarrow y \leftrightarrow z)\, \Gamma(y \leftarrow z)\, dz. \tag{7}$$

Proof The adjoint operator "*" is a linear operator, and the identity operator $\mathcal{I}$ is self-adjoint, so $(\mathcal{I} - \mathcal{T})^* = \mathcal{I} - \mathcal{T}^*$. The proof then follows immediately from the previous lemma. □

4 Relationship of radiance and directional importance

The transport equation for directional importance does not in itself impart any particular units on this new quantity Γ. In order to give meaningful units to directional importance, we make the following observations.

We define a new quantity, *incoming radiance* $\mathrm{L}^{\leftarrow}(y \leftarrow x)$, by the following identity:

$$\mathrm{L}^{\leftarrow}(y \leftarrow x) \;\equiv\; G(x \leftrightarrow y)\, \mathrm{L}(x \to y). \tag{8}$$

Intuitively, $\mathrm{L}^{\leftarrow}$ is the arriving radiance at y from the direction of x, per unit area at x and y. Thus, incoming radiance $\mathrm{L}^{\leftarrow}$ is measured in [watt $\cdot$ meter^{-4}]; it describes power (energy flux) per unit area of receiver per unit area of source.

We can also write an equation for (outgoing) radiance $\mathrm{L}(y \to z)$ as a function of incoming radiance at y:

$$\mathrm{L}(y \to z) \;= \mathring{\mathrm{L}}\,(y \to z) + \int_{x} f_r(x \leftrightarrow y \leftrightarrow z)\, \mathrm{L}^{\leftarrow}(y \leftarrow x)\, dx. \tag{9}$$

This equation is easily checked by plugging in the definition from equation (8).

Finally, multiplying through by $G(z \leftrightarrow y)$, we can derive a transport equation for incoming radiance $\mathrm{L}^{\leftarrow}$:

$$\mathrm{L}^{\leftarrow}(z \leftarrow y) \;=\; \mathring{\mathrm{L}}^{\leftarrow}(z \leftarrow y) + G(z \leftrightarrow y) \int_x f_r(x \leftrightarrow y \leftrightarrow z)\, \mathrm{L}^{\leftarrow}(y \leftarrow x)\, dx, \quad (10)$$

where

$$\mathring{\mathrm{L}}^{\leftarrow}(z \leftarrow y) \;\equiv\; G(y \leftrightarrow z)\, \mathring{\mathrm{L}}\,(y \rightarrow z).$$

Note that this equation is identical to Kajiya's "rendering equation" [4], and that "incoming radiance" is exactly the same as Kajiya's "two point transport intensity."

Furthermore, note that equation (10) differs from the transport equation for directional importance (7) only in the order of the arguments to the BRDF; however, this function is symmetric by reciprocity, so the two transport equations are in fact identical:

Theorem *Incoming radiance* $\mathrm{L}^{\leftarrow}$ *and directional importance* Γ *satisfy the same transport equation.*

In an entirely similar manner, we can also define *outgoing directional importance* $\Gamma^{\rightarrow}(x \rightarrow y)$ by the relation

$$\Gamma(y \leftarrow x) \;=\; G(x \leftrightarrow y)\Gamma^{\rightarrow}(x \rightarrow y). \quad (11)$$

Intuitively, $\Gamma^{\rightarrow}$ is the importance leaving x in the direction of y, whereas Γ is the importance arriving at y from the direction of x.

The following theorem is then easy to check:

Theorem *Radiance* L *and outgoing directional importance* $\Gamma^{\rightarrow}$ *satisfy the same transport equation.*

It is therefore natural to give outgoing directional importance the same units as radiance, and to propagate outgoing directional importance from the eye in exactly the same fashion as radiance is propagated from the light sources.

Since the two formulations—incoming or outgoing—are mathematically equivalent, there would seem at first to be no particular advantage of using one formulation over the other for computing the transport of radiance and directional importance.

However, inspection of equations (1) and (10) reveals one major difference: if we ignore for the moment any ideal specular reflection, then the outgoing quantities at any surface point are continuous functions of angle, whereas the incoming quantities may have discontinuities, occuring, for instance, along shadow boundaries. Such discontinuities would hinder the use of spherical harmonics to represent the distributions [7] since spherical harmonic expansions of discontinuous functions converge slowly.

By using the outgoing formulations of both radiance and directional importance, we can employ spherical harmonics to represent both types of quantities using relatively few terms. Using outgoing formulations for both quantities also simplifies the development of an importance-driven radiance algorithm, since it allows both radiance and importance to be transported identically.

5 Preliminary results

Smits *et al.* [8] demonstrated that the use of importance can dramatically reduce the time to compute radiosities when much of the scene is invisible. Preliminary results, based on a hierarchical algorithm we are currently developing, indicate that similar speedups are possible when transporting radiance and directional importance. (The details of our algorithm and our methods for estimating error will be presented in a subsequent paper.)

In addition, our results indicate that directional importance allows speedups even when all objects are visible. Consider, for example, the four different "Cornell box" scenes in figure 1. Image (a) shows the standard Cornell box, with a diffusely-emitting overhead light source and diffusely-reflecting walls. Image (b) shows the same diffuse box, with diffuse lights placed near the walls; here, the radiance fall-off with distance is accentuated. In image (c) the walls are again diffuse reflectors, but the light sources are highly directional—very little light is emitted toward the eye. In image (d), the light sources are directional and the walls are glossy.

We have used flat-shaded versions of images (b) and (d) as reference solutions, as shown in figure 2. These reference images were generated using our radiance program, but with importance disabled. For each image, the program was allowed to refine the solution until the root-mean-squared discrepancy between successive images was less than 1%.

The solutions were computed with and without importance, and compared to the reference solutions. Figure 3 plots root-mean-squared error versus the number of links in the hierarchical system for images (b) and (d). Note that the importance-driven algorithm requires twice as many links as the algorithm without importance for the early iterations since importance is propagated in addition to radiance. However, as refinement proceeds, the overhead of transporting importance is eventually repayed by a reduction in the number of links.

The glossy scene of image (d) demonstrates the benefit of using directional importance even when all objects are visible. This scene has four light sources, but only one of the lights contributes radiance in the direction of the eye. In this case, for a 1% error, the importance-driven algorithm requires 1/4 the number of radiance links, along with roughly an equal number of importance links, yielding 1/2 the number of links overall. Note, however, that the gain from using importance can be made arbitrarily high by increasing the number of light sources that do not contribute visibly to the image, even if all objects are visible. Of course, when objects are also hidden, the gain is even greater.

Note also that the use of directional importance provides moderate gains even when all objects are visible in purely diffuse environments, as indicated by the results for image (b). This slight advantage occurs because our importance-driven refinement strategy is based on *projected* area, which depends on the particular orientation of each patch with respect to the viewer. Patches with smaller projected areas receive less importance since they contribute less to the visible scene. Thus, as the error tolerance is reduced, the importance-driven algorithm eventually reduces the number of links on obliquely oriented patches enough to produce an overall speedup.

6 Summary

Using ideas from linear operator theory, we have developed a theoretical framework for extending the definition of importance to the general case of directional reflection. The framework was also used to show that a variant of importance, namely outgoing directional importance, satisfies the same transport equation as radiance. The distribution of outgoing directional importance in a scene is therefore equivalent to the distribution of radiance that would occur if the eye were the only radiant emitter.

We are currently developing a hierarchical algorithm based on outgoing directional importance. Although numerous aspects of the algorithm require further investigation, our preliminary results indicate that this adjoint formulation has all the advantages of importance as used in Smits *et al.* [8]. Additionally, directional importance also seems to provide speedups even when all objects are visible. We plan to offer a detailed description of the algorithm in a subsequent paper.

References

[1] Robert L. Cook, Thomas Porter, and Loren Carpenter. Distributed ray tracing. *Computer Graphics*, 18(3):137–145, July 1984.

[2] Pat Hanrahan, David Salzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. *Computer Graphics*, 25(4):197–206, July 1991.

[3] David S. Immel, Michael F. Cohen, and Donald P. Greenberg. A radiosity method for non-diffuse environments. *Computer Graphics*, 20(4):133–142, August 1986.

[4] James T. Kajiya. The rendering equation. *Computer Graphics*, 20(4):143–150, August 1986.

[5] S. N. Pattanaik and S. P. Mudur. The potential equation and importance in illumination computations. In *Computer Graphics Forum*, 1993. To appear.

[6] Min-Zhi Shao, Qun-Sheng Peng, and You-Dong Liang. A new radiosity approach by procedural refinements for realistic image syntehsis. *Computer Graphics*, 22(4):93–102, August 1988.

[7] François Sillion, James R. Arvo, Stephen H. Westin, and Donald P. Greenberg. A global illumination solution for general reflectance distributions. *Computer Graphics*, 25(4):187–196, July 1991.

[8] Brian E. Smits, James R. Arvo, and David H. Salesin. An importance-driven radiosity algorithm. *Computer Graphics*, 26(2):273–282, July 1992.

[9] Gregory J. Ward, Francis M. Rubinstein, and Robert D. Clear. A ray tracing solution for diffuse interreflection. *Computer Graphics*, 22(4):85–92, August 1988.

Figure 1 Four Gouraud-shaded Cornell boxes with different illumination and reflectances: (a) diffuse overhead light source, diffuse reflection; (b) diffuse light sources, diffuse reflection; (c) directional light sources, diffuse reflection; (d) directional light sources, glossy reflection.

Figure 2 Flat-shaded Cornell boxes, corresponding to images (b) and (d) of figure 1.

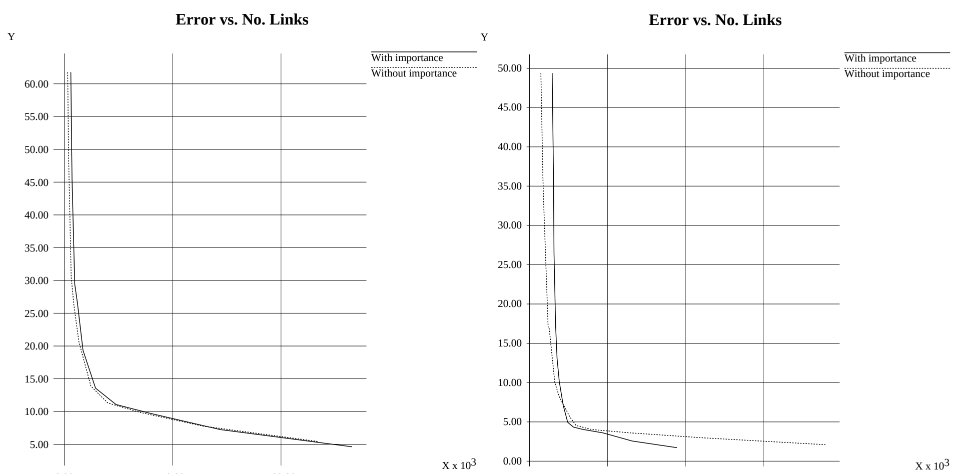

Figure 3 Root-mean-squared error versus number of links for images (b) and (d) of figure 1.