Wavelet Radiative Transfer and Surface Interaction

Robert R. Lewis
bobl@cs.ubc.ca

Imager Computer Graphics Lab
University of British Columbia
Department of Computer Science

20 February, 1996

1 Introduction

Illumination is the study of how light interacts with matter to produce visible scenes. In computer graphics, we use illumination to produce "realistic" images.

Local illumination describes the interaction of light with a single, small volume or surface element with given incident and viewing directions. Figure 1 shows the typical geometry and and nomenclature for local illumination studies.

Global illumination describes how light is distributed in a *scene*: a collection of objects, including light sources, immersed in a given medium. Global illumination solutions must consider multiple reflections. Global illumination solutions are built on top of local illumination solutions.

In this paper, we will advance a new approach to an illumination solution that is intermediate between local and global illumination. Using wavelets, we are able to treat the interaction between two surfaces and the interaction of a surface with a radiation field in a source-to-destination model that applies to whole surfaces, not just small elements. We are continuing work to extend this to a fully global solution.

Wavelets are relatively recent additions to the rendering toolkit. They were first used by Gortler et al. [gort93] to solve the radiosity equations. Schröder et al. [schr93] and Christensen et al. [chri94] applied them to non-diffuse situations. What we present here may be considered a further development of that work.

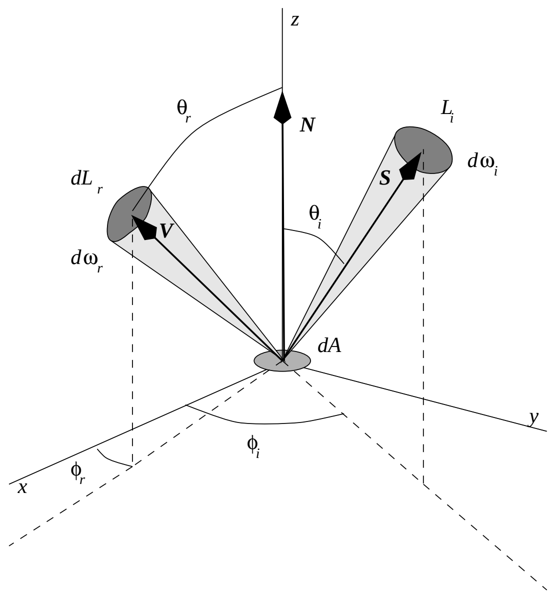

Figure 1: Geometry

2 Radiative Transfer and Surface Interaction

Let us first discuss some of the basics of how light is represented. The fundamental quantity is *radiance*, the amount of power passing in a given direction though a given surface per unit area (perpendicular to the direction of travel) per unit solid angle[1].

[1] In this report, we will take radiance to be monochromatic and assume we can construct a polychromatic images by combining independently computed monochromatic images. We

Radiance at a point $\mathbf{P}$ in a direction $\mathbf{S}$ is usually represented by a function $L(\mathbf{P}, \mathbf{S})$.

Radiance's most useful property is its invariance: In a non-participating medium, the radiance given off at a point $\mathbf{P_o}$ on a surface in a direction $\mathbf{S}$ is constant until reaching another surface. We can express this as the action of a transport operator $\mathcal{T}$:

$$L(\mathbf{P}, \mathbf{S}) = \mathcal{T} L = L(\mathbf{P}_o(\mathbf{P}, \mathbf{S}), \mathbf{S}) \tag{1}$$

This principle underlies raytracing.

We take the fundamental equation describing surface interaction to be (cf. Foley, et al. [fole90] or Cohen and Wallace [cohe93]):

$$\begin{aligned} L(\mathbf{V}) &= L_e(\mathbf{V}) \\ &+ \int_{\Omega_\mathbf{N}^R} f_r(\mathbf{S}^{+\prime}, \mathbf{V})\, L_i(\mathbf{S}^{+\prime})\, \left|\mathbf{N} \cdot \mathbf{S}^{+\prime}\right|\, d\omega_i' \\ &+ \int_{\Omega_\mathbf{N}^T} f_t(\mathbf{S}^{-\prime}, \mathbf{V})\, L_i(\mathbf{S}^{-\prime})\, \left|\mathbf{N} \cdot \mathbf{S}^{-\prime}\right|\, d\omega_i' \end{aligned} \tag{2}$$

where L is the total radiance given off of a surface with normal $\mathbf{N}$, L_e is the surface emissivity, L_i is the incident radiance, $\Omega_\mathbf{N}^R$ is the reflection hemisphere (contains N, the "viewing" direction $\mathbf{V}$, and the "positive source" direction $\mathbf{S}^+$), $\Omega_\mathbf{N}^T$ is the transmission hemisphere (opposite $\Omega_\mathbf{N}^R$, containing the "negative source" direction $\mathbf{S}^-$), f_r is the bidirectional reflectance distribution function (BRDF), and f_t is the bidirectional transmittance distribution function (BTDF). Note that we use the prime ("$\prime$") to indicate bound variables of integration and, since it symbolically factors out of (2), we ignore the spatial variation of the radiances and distribution functions.

We represent combined reflection and transmission surface interaction by the integral operator $\mathcal{S}$. We can treat the emissivity term separately.

2.1 Radiance in Nusselt Coordinates

If we confine our discussion to flat surfaces, we can assume a parameterization for $\mathbf{P}$ of (x, y). $\mathbf{S}$ is then typically represented in polar and azimuthal coordinates (θ, ϕ) according to the local frame of reference.

Consider the x, y, and z direction cosines corresponding to a direction (θ, ϕ):

$$\mu_x = \sin\theta\, \cos\phi \quad \mu_y = \sin\theta\, \sin\phi \quad \mu_z = \cos\theta \tag{3}$$

also ignore polarization.

For reasons that will be obvious later, let us use these to make a change of the directional variables from (θ, ϕ) to (κ, λ):

$$\kappa = \tfrac{\mu_x + 1}{2} \quad \lambda = \tfrac{\mu_y + 1}{2} \tag{4}$$

To convert integration over (θ, ϕ) to integration over (κ, λ), the determinant of the Jacobian is:

$$\left|\frac{\partial(\theta, \phi)}{\partial(\kappa, \lambda)}\right| = \frac{4}{|\cos\theta \sin\theta|} \tag{5}$$

so, assuming L_i is zero for directions outside the unit circle, (2) becomes

$$\begin{aligned} L &= L_e + 4 \iint_0^1 \left[f_r(\mathbf{S}^{+\prime}, \mathbf{V})\, L_i(\mathbf{S}^{+\prime}) \right. \\ &\qquad \left. + f_t(\mathbf{S}^{-\prime}, \mathbf{V})\, L_i(\mathbf{S}^{-\prime}) \right] d\kappa_i'\, d\lambda_i' \end{aligned} \tag{6}$$

That the integral no longer contains trigonometric functions should come as no surprise. We have simply used a differential form of the "Nusselt analog" ([nuss28], but see Cohen and Wallace [cohe93] for a description in English): The amount of power per unit area transferred from a differential solid angle $d\omega_i$ is proportional to the area of the surface that the projection of $d\omega_i$ on a unit sphere subtends, $d\kappa_i\, d\lambda_i$. For this reason, we refer to κ and λ as "Nusselt coordinates".

We also note that since, $\mu_x^2 + \mu_y^2 + \mu_z^2 = 1$, we can express $\mathbf{S}^+$, $\mathbf{S}^-$, and $\mathbf{V}$ all unabiguously in terms of their respective incident and reflected κ's and λ's, since each vector is defined only over a hemisphere, not the whole directional sphere. Simply put, it's always clear which sign to attach to the square root.

2.2 Radiance Representation

What are the characteristics of a four-dimensional radiance distribution $L(x, y, \kappa, \lambda)$? The easiest way to visualize this is as "light through a window" where the position of an observer on a window is (x, y) and he or she is looking in direction (κ, λ). For a fixed direction, the resulting two-dimensional projection is a parallel projection. (The special case $(\kappa, \lambda) = (\frac{1}{2}, \frac{1}{2})$ is an orthographic projection.) For a fixed position, the distribution in (κ, λ) would be a "fisheye" view.

In both cases, what the observer sees is an image, so we can infer that dealing with radiance distributions should be like dealing with images.

Radiance $L(x, y, \kappa, \lambda)$ at a point on a surface is a potentially discontinuous, generally nonanalytic function. We can approximate it with a finite element expansion with N_f degrees of freedom:

$$L(x, y, \kappa, \lambda) = \sum_{j=1}^{N_f} b_j B_j(x, y, \kappa, \lambda) \tag{7}$$

Choices for the basis functions B_i include box discretization (*a la* Fournier et al.'s FIAT [four90]), Fourier, discrete cosine, orthogonal polynomials, and wavelets. We are particularly interested in wavelets because, unlike the other bases listed, their basis functions are of limited support and they can represent discontinuities compactly. They are also capable of considerable compression.

3 Wavelet Properties

In this section, we will review some of the properties of wavelets that make them particularly suitable for the representation of radiance. The standard reference on wavelets is Daubechies [daub92], from which much of this section is derived.

3.1 One-Dimensional Wavelets

Before proceeding to multidimensional wavelets, we should first cover several important properties of one-dimensional wavelets.

3.1.1 Scaling Functions and Wavelets

Wavelets are built from *scaling functions*, which we define by enumerated dilations and translations of a base scaling function $\phi(x)$ of the form:

$$\phi_{lm}(x) = 2^{l/2}\phi(2^l x - m) \tag{8}$$

each level l corresponds to a function space V_l, which is part of a nested sequence of subspaces $\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \ldots$. Scaling functions have the property that

$$\int_{-\infty}^{+\infty} \phi(x)dx \neq 0 \tag{9}$$

We define a wavelet function space W_l as composed of those functions that need to be added to a given space V_l to span the next finer space V_{l+1}: $V_{l+1} = V_l \oplus W_l$. The basis functions for W_l are then dilations and translations of a *mother wavelet* $\psi(x)$:

$$\psi_{lm}(x) = 2^{l/2}\psi(2^l x - m) \tag{10}$$

Wavelets have the property

$$\int_{-\infty}^{+\infty} \psi(x)dx = 0 \tag{11}$$

Figures 2-4 show some commonly-used wavelets and their corresponding smoothing functions.

3.1.2 Multiresolution Refinement Equations

Since $\phi(x) \in V_0$ and $V_0 \subset V_1$, we can write $\phi(x)$ as a linear combination of the basis functions $\phi(2x - j)$ for V_1:

$$\phi(x) = \sqrt{2}\sum_j h_j \phi(2x - j) \tag{12}$$

This also holds for ψ:

$$\psi(x) = \sqrt{2}\sum_j g_j \phi(2x - j) \tag{13}$$

These are the *dilation* or *refinement equations.* Wavelet bases differ principally in their choices of $\{h_j\}$ (which determines $\{g_j\}$).

Using the enumerated bases:

$$\begin{aligned} \phi_{lm}(x) &= \sqrt{2}\sum_j h_j \phi_{l+1,2m+j}(x) \\ \psi_{lm}(x) &= \sqrt{2}\sum_j g_j \phi_{l+1,2m+j}(x) \end{aligned} \tag{14}$$

3.1.3 Orthogonal Wavelets

We define the inner product of two functions f and g with respect to x[2]

$$\langle f \mid g \rangle_x \equiv \int_{-\infty}^{+\infty} f(x')g(x')dx' \tag{15}$$

Some wavelets are orthogonal:

$$\langle \psi_{lm} \mid \psi_{l'm'} \rangle_x = \delta_{ll'}\delta_{mm'} \tag{16}$$

[2] Using x as a subscript is non-traditional, but will become useful when we speak of multidimensional wavelets.

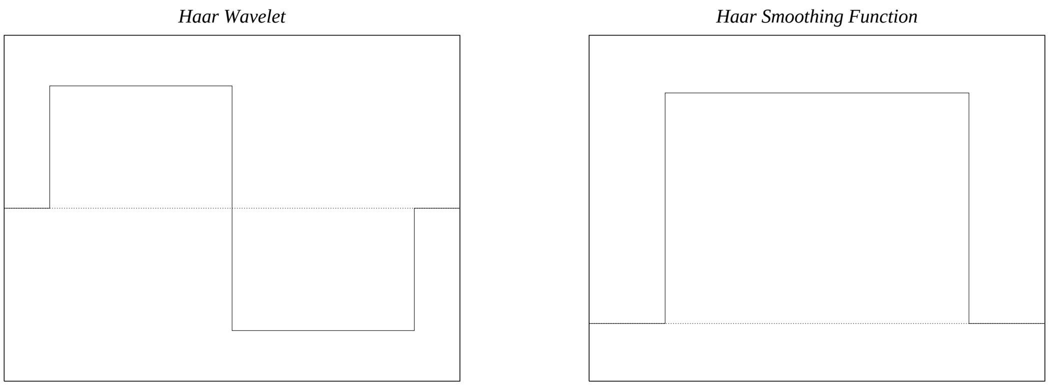

Figure 2: Haar Wavelet and Smoothing Functions. Haar wavelets are the simplest possible wavelet and are the most commonly used wavelets in graphics. They have only one vanishing moment.

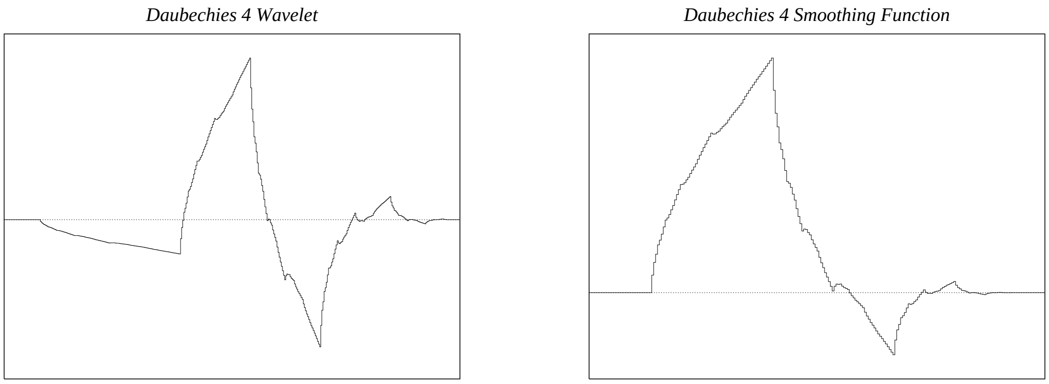

Figure 3: Daubechies-4 Wavelet and Smoothing Functions. Daubechies wavelets were the first compact orthogonal wavelets discovered. This particular wavelet has two vanishing moments.

But these have undesirable features. First, except for Haar, they aren't symmetric. Second, they don't include useful functions like splines.

3.1.4 Biorthogonal Wavelets

We can construct biorthogonal bases by using four functions instead of two: wavelets ϕ_{lm} and $\tilde{\phi}_{lm}$ and smoothing functions ψ_{lm} and $\tilde{\psi}_{lm}$. These are defined defined, respectively, by four sets of coefficents: $\{h_j\}$, $\{\tilde{h}_j\}$, $\{g_j\}$, and $\{\tilde{g}_j\}$.

For these,

$$\left\langle \psi_{lm} \mid \tilde{\psi}_{l'm'} \right\rangle_x = \delta_{ll'} \delta_{mm'} \tag{17}$$

In the rest of this section, we'll assume the more general biorthogonality, since we can always treat orthogonal wavelets as a special case of biorthogonal ones.

3.1.5 Wavelet Projections and Approximation

Let us discuss the ability of a wavelet representation to approximate an arbitrary function f. Let $\mathcal{P}_l f$ be the projection of a function $f \in L^2$ into the subspace V_l:

$$\mathcal{P}_l f(x) = \sum_m \left\langle f \mid \tilde{\phi}_{lm} \right\rangle_x \phi_{lm}(x) \tag{18}$$

It can be shown

$$\| f - \mathcal{P}_l f \|_2 \leq C 2^{-lN_v} \sqrt{\sum_n \| f^{(n)} \|^2} \tag{19}$$

where N_v is the number of vanishing moments of the

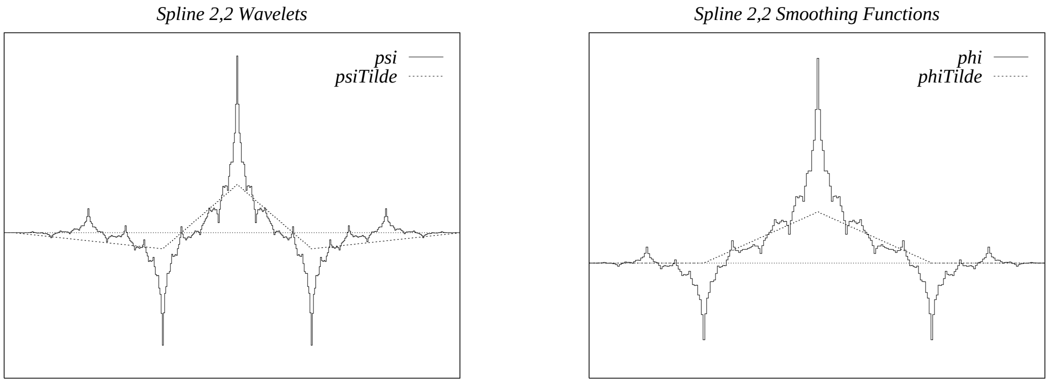

Figure 4: Linear Spline Wavelet And Smoothing Functions. These wavelets are biorthogonal and have two vanishing moments. Unfortunately, while $\tilde{\phi}$ and $\tilde{\psi}$ have a simple closed form, this is not true for ϕ and ψ.

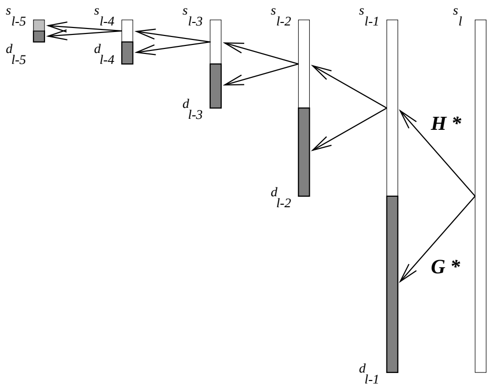

Figure 5: The Fast Wavelet Transform

wavelets, i. e. for $n = 0, \ldots, N_v - 1$

$$\int x^n \tilde{\psi}(x) dx = \left\langle x^n \mid \tilde{\psi}_{00} \right\rangle_x = 0 \qquad (20)$$

What (19) means is that we can always decrease the L^2 error of a wavelet approximation by increasing the number of levels l or by choosing wavelets with a higher number of vanishing moments.

3.1.6 The Fast Wavelet Transform

Given a set of data $s_{lm}, m = 0 \ldots N$, where $N = 2^l - 1$, we treat these as coefficients of ϕ_{lm} and can compute a wavelet transform in $O(N)$ operations, as shown schematically in Figure 5. This is another advantage of wavelet methods over Fourier techniques, which typically require $O(N \log N)$ operations. The transform leaves us with $2^l - 1$ wavelet coefficients (dark gray in Figure 5) and 1 smoothing coefficient (light gray). The hierarchical arrangement of these coefficients is sometimes referred to as the "wavelet pyramid".

3.1.7 Wavelet Compression

Even with the wavelet pyramid, we've still got N coefficients, so we haven't saved any storage (yet).

It can be shown that if f is of the form

$$f(x) = \sum_{l,m} w_{lm} \psi_{lm}(x) \qquad (21)$$

and we approximate it by

$$F(x) = \sum_{w_{lm} \in \tilde{W}} w_{lm} \psi_{lm}(x) \qquad (22)$$

where the $\tilde{W}$ is a subset of the set of coefficents $\{w_{lm}\}$, that

$$\| f - F \|^2 = \sum_{w_{lm} \notin \tilde{W}} |w_{lm}|^2 \qquad (23)$$

This gives us a convenient error metric. It also tells us that the optimal compression scheme discards the coefficients with smaller absolute magnitudes first – a thresholding process.

3.2 Multidimensional Wavelets

For D-dimensional coordinates

$$\mathbf{q} = (q_1, q_2, \ldots, q_D), \tag{24}$$

we can define a set of multidimensional wavelet basis functions indexed by a *multiresolution index*

$$\mathbf{j} = (\nu^j, l_1^j, m_1^j, l_2^j, m_2^j, \ldots, l_D^j, m_D^j): \tag{25}$$

where ν^j determines the combination of one dimensional smoothing and wavelet functions:

$$B_{\mathbf{j}}(\mathbf{q}) = \begin{cases} \phi_{l_1^j m_1^j}(q_1)\phi_{l_2^j m_2^j}(q_2)\ldots\phi_{l_D^j m_D^j}(q_D) & \nu^j = 0 \\ \psi_{l_1^j m_1^j}(q_1)\phi_{l_2^j m_2^j}(q_2)\ldots\phi_{l_D^j m_D^j}(q_D) & \nu^j = 1 \\ \phi_{l_1^j m_1^j}(q_1)\psi_{l_2^j m_2^j}(q_2)\ldots\phi_{l_D^j m_D^j}(q_D) & \nu^j = 2 \\ \vdots & \\ \psi_{l_1^j m_1^j}(q_1)\psi_{l_2^j m_2^j}(q_2)\ldots\psi_{l_D^j m_D^j}(q_D) & \nu^j = 2^D - 1 \end{cases} \tag{26}$$

We refer to the special case of $\nu_j = 0$ as the "pure smoothing" component, as the corresponding basis function is made up of only smoothing functions.

We can apply (17) multidimensionally. If we have $\mathbf{j}$ as in (25) and define $\mathbf{k}$ as

$$\mathbf{k} = (\nu^k, l_1^k, m_1^k, l_2^k, m_2^k, \ldots, l_D^k, m_D^k) \tag{27}$$

then

$$\left\langle \tilde{B}_{\mathbf{j}} \mid B_{\mathbf{k}} \right\rangle_{\mathbf{q}} = \delta_{\nu^j \nu^k} \prod_{d=1}^{D} \delta_{l_d^j l_d^k} \delta_{m_d^j m_d^k} \tag{28}$$

This is the "standard" multidimentsional Cartesian product basis. It is also possible to constrain $l_1^j = l_2^j = \ldots = l_D^j \equiv l^j$, resulting in the so-called "nonstandard" basis. In general multidimensional (especially image-oriented) applications, as cited in Daubechies [daub92] and in Schröder et al. [schr93], the nonstandard bases are preferred because of their "square" (hypercubic, in our case) support.

In this paper, we are considering 4-dimensional, nonstandard basis functions, so let us enumerate the coordinates with the 4-vector $\mathbf{q} = (u, v, \kappa, \lambda)$ and the basis functions with $(\nu, l, m_u, m_v, m_\kappa, m_\lambda)$.

Using (8), (10), (12), (13) and (26), all the basis functions at level l of the pyramid can be written in terms of the $\nu = 0$ basis functions at level $l + 1$:

$$B_{\nu l \mathbf{m}}(\mathbf{q}) = 4 \begin{cases} \sum_{\mathbf{m}'} h_{m'_u} h_{m'_v} h_{m'_\kappa} h_{m'_\lambda} B_{0(l+1)(2\mathbf{m}+\mathbf{m}')}(\mathbf{q}) & \nu = 0 \\ \sum_{\mathbf{m}'} g_{m'_u} h_{m'_v} h_{m'_\kappa} h_{m'_\lambda} B_{0(l+1)(2\mathbf{m}+\mathbf{m}')}(\mathbf{q}) & \nu = 1 \\ \sum_{\mathbf{m}'} h_{m'_u} g_{m'_v} h_{m'_\kappa} h_{m'_\lambda} B_{0(l+1)(2\mathbf{m}+\mathbf{m}')}(\mathbf{q}) & \nu = 2 \\ \vdots & \\ \sum_{\mathbf{m}'} g_{m'_u} g_{m'_v} g_{m'_\kappa} g_{m'_\lambda} B_{0(l+1)(2\mathbf{m}+\mathbf{m}')}(\mathbf{q}) & \nu = 15 \end{cases} \tag{29}$$

where $\mathbf{m}' \equiv (m'_u, m'_v, m'_\kappa, m'_\lambda)$.

So for any function f,

$$\langle f \mid B_{\nu l \mathbf{m}} \rangle_q = \sum_{\mathbf{m}'} W_{\nu \mathbf{m}'} \left\langle f \mid B_{0(l+1)(2\mathbf{m}+\mathbf{m}')} \right\rangle_q \tag{30}$$

where $W_{\nu \mathbf{m}'}$ is (4 times) a product of smoothing and wavelet coefficients.

4 Wavelet Radiance Properties

Apart from compression, representing radiance in terms of a wavelet basis with direction expressed in Nusselt coordinates makes several calculations of relevance to illumination computation easier. Notice that these all act directly on wavelet coefficients themselves and do not require an inverse wavelet transform.

4.1 Irradiance Computation

Irradiance is defined as

$$\begin{aligned} E(x, y) &= \int_{\Omega_{\mathbf{N}}^R} L_i(x, y, \theta', \phi') \, |\mathbf{N} \cdot \mathbf{S}'| \, d\omega'_i \\ &= 4 \iint_0^1 L_i(x, y, \kappa', \lambda') \, d\kappa'_i \, d\lambda'_i \end{aligned} \tag{31}$$

But if we have

$$L_i(x, y, \kappa, \lambda) = \sum_{\mathbf{j}} b_{\mathbf{j}} B_{\mathbf{j}}(x, y, \kappa, \lambda) \tag{32}$$

then

$$E(x, y) = 4 \sum_{\mathbf{j}} b_{\mathbf{j}} \langle B_{\mathbf{j}} \mid 1 \rangle_{\kappa, \lambda} \tag{33}$$

is the wavelet representation of the irradiance. The inner product on the right hand side is usually easy to compute analytically, making particular use of (11) to eliminate some of the non-pure smoothing coefficients[3].

4.2 Transport

We represent radiance as

$$L(\mathbf{q}) = \sum_{\mathbf{k}} b_{\mathbf{k}} B_{\mathbf{k}}(\mathbf{q}) \tag{34}$$

where

$$b_{\mathbf{k}} = \left\langle L \mid \tilde{B}_{\mathbf{k}} \right\rangle_{\mathbf{q}} \tag{35}$$

and $\mathbf{k}$ is defined as in (27).

Radiance travels from a source point $\mathbf{q_s}$ to a destination point $\mathbf{q_d}$. If we have a mapping of $\mathbf{q_s} \to \mathbf{q_d}$, we can compute

$$L_d(\mathbf{q_d}) = \sum_{\mathbf{k}} b^d_{\mathbf{k}} B_{\mathbf{k}}(\mathbf{q_d}) \tag{36}$$

where

$$\begin{aligned} b^d_{\mathbf{k}} &= \left\langle L_s(q_s(\cdot)) \mid \tilde{B}_{\mathbf{k}} \right\rangle_{q_d} \\ &= \sum_{\mathbf{j}} b^s_{\mathbf{j}} T_{\mathbf{kk}}, \end{aligned} \tag{37}$$

$\mathbf{j}$ is defined as in (25), and we define geometry-dependent "transport coefficents"

$$T_{\mathbf{jk}} \equiv \left\langle B_{\mathbf{j}}(q_s(\cdot)) \mid \tilde{B}_{\mathbf{k}} \right\rangle_{\mathbf{q_d}}. \tag{38}$$

$T_{\mathbf{jk}}$ is large, but $b^s_{\mathbf{j}}$ is sparse. In some applications, it might be possible to compute coupling coefficients as needed.

Using the multidimensional refinement shown in (29), given $T_{(0l\mathbf{m})\mathbf{k}}$ on level l, we can compute all coefficients on the level above it:

$$T_{(\nu(l-1)\mathbf{m})\mathbf{k}} = \sum_{\mathbf{m}'} W_{\nu\mathbf{m}'} T_{0l(2\mathbf{m}+\mathbf{m}')\mathbf{k}} \tag{39}$$

and given $T_{\mathbf{j}(0l\mathbf{m})}$ on level l, we can compute

$$T_{\mathbf{j}(\nu(l-1)\mathbf{m})} = \sum_{\mathbf{m}'} W_{\nu\mathbf{m}'} T_{\mathbf{j}(0l(2\mathbf{m}+\mathbf{m}'))} \tag{40}$$

where

$$\mathbf{m} = (m_u, m_v, m_\kappa, m_\lambda). \tag{41}$$

This means that we can compute all transport coefficients strictly in terms of pure smoothing components.

[3] It's important to remember that the inner product is taken over $[0,1]$, while (11) is over $(-\infty, \infty)$.

4.3 Surface Interaction

If we represent the BRDF in Nusselt coordinates with the *dual* wavelet basis:

$$f_r(\kappa_s, \lambda_s, \kappa_r, \lambda_r) = \sum_{\mathbf{j}} f_{\mathbf{j}} \tilde{B}_{\mathbf{j}}(\kappa_s, \lambda_s, \kappa_r, \lambda_r) \tag{42}$$

and

$$L_i(x, y, \kappa_s, \lambda_s) = \sum_{\mathbf{k}} b_{\mathbf{k}} B_{\mathbf{k}}(x, y, \kappa_s, \lambda_s) \tag{43}$$

where

$$\begin{aligned} \mathbf{j} &= (\nu_j, l_j, m^i_\kappa, m^i_\lambda, m^j_\kappa, m^j_\lambda) \\ \mathbf{k} &= (\nu_k, l_k, m^k_u, m^k_v, m^k_\kappa, m^k_\lambda) \end{aligned} \tag{44}$$

so, applying (6), the reflected radiance is

$$\begin{aligned} L_r &= 4 \int\int_0^1 f_r(\mathbf{S}^{+\prime}, \mathbf{V})\, L_i(\mathbf{S}^{+\prime}) d\kappa'_s\, d\lambda'_s \\ &= 4 \sum_{\mathbf{j}} \sum_{\mathbf{k}} f_{\mathbf{j}} b_{\mathbf{k}} \left\langle \tilde{B}_{\mathbf{r}} \mid B_{\mathbf{i}} \right\rangle_{\kappa_s, \lambda_s} \end{aligned} \tag{45}$$

Making use of biorthogonality via (28), we derive a wavelet representation of the post-interaction radiance:

$$L_r = \sum_{\mathbf{n}} b^r_{\mathbf{n}} B_{\mathbf{n}}(x, y, \kappa_s, \lambda_s) \tag{46}$$

where

$$\begin{aligned} b^r_{\mathbf{n}} &= b^r_{\nu l m_u m_v m_\kappa m_\lambda} \\ &= 4 \sum_{m^i_\kappa, m^i_\lambda} f_{\nu l m^i_\kappa m^i_\lambda m_\kappa, m_\lambda} b^i_{\nu l m_u m_v m^i_\kappa m^i_\lambda}. \end{aligned} \tag{47}$$

If both $\mathbf{f}$ and $\mathbf{b}$ have sparse representations, $\mathbf{b^r}$ can be efficiently calculated by traversing their lookup tables.

We can do the same thing with a BTDF and so represent general surface interactions: reflection, refraction, and transmission.

5 Implementation

Let's apply some of the concepts of the previous section to a classic illumination problem: the transport of radiation between two arbitrarily-oriented quadrilaterals.

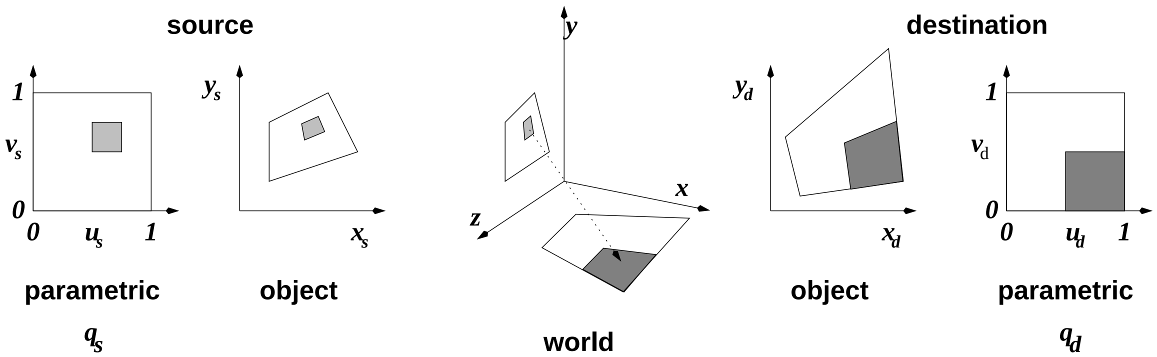

Figure 6: Quadrilateral Coordinate Systems

5.1 Coordinate Systems

We must deal with several coordinate systems, as shown in Figure 6. So far, we have dealt with a positional distribution of radiance defined on a unit square. In treating quadrilaterals, we still take radiances to be represented on on the unit square, but we now include a bilinear transformation to take us from the (now) parametric unit square to an arbitrary quadrilateral[4] and back. The quadrilaterals are each defined in their respective 2-dimensional "object coordinates". Furthermore, to establish the $\mathbf{q_s} \rightarrow \mathbf{q_d}$ mapping we allow a conventional affine mapping of each quadrilateral to a "world" coordinate system.

5.2 Choice of Wavelet

Our discussion Sections 3 and 4 did not depend on any particular wavelet. For implementation purposes, we have to choose one. There is an extremely good reason for choosing Haar wavelets, at least for a first attempt.

As we mentioned in Section 3.1.6, the fast wavelet transform can be performed in $O(N)$ time, but if we allow for a varying dimensionality D and wavelet basis, it is easy to see from (29) that the efficiency is actually $O(|\mathbf{h}|^D_{\max} N)$ where $|\mathbf{h}|_{\max}$ is the maximum width of the $\mathbf{h}$ and $\tilde{\mathbf{h}}$ filters.

For this reason, as the dimensionality increases, the rapidly-increasing operation count makes narrower filters increasingly desirable, even though wider filters have better approximation properties, in general. Since Haar is the narrowest possible wavelet filter ($|\mathbf{h}|_{\max} = 2$), it seems a wise strategy to make any multidimensional efforts first with Haar and move to wider bases later.

An additional advantage of Haar wavelets over the others is the simplification of the calculation of the transport coefficents. As (39) and (40) have shown, these coefficients can be computed entirely in terms of pure smoothing calculations. A four-dimensional Haar pure smoothing basis is a function that is constant within a hypercube and zero outside of it. The resulting transport coefficients are volume integrals of the overlap between such a hypercube in destination parametric space and the object which is a projection of a hypercube in source parametric space into the destination space[5].

5.3 Transport Coefficient Computation

As the preceding section suggests, using Haar wavelets turns transport coefficient computation into volume integral evaluation. There are two practical problems that then arise from the computation of that volume. (Throughout this discussion, we will be speaking in general terms.)

5.3.1 Some Source Points Don't Project Into Destination Space

The source hypercube defines a range of positional and directional coordinates $\mathbf{q_s}$. Not all of these coordinates may map to points in the destination plane,

[4] We allow triangles as a degenerate case of quadrilaterals.

[5] This may not be such a great simplification. After all, any arbitrarily complex 3-dimensional integration may be trivially turned into a 4-dimensional volume integral!

much less the destination quadrilateral. This complicates any attempt at direct evaluation of the transport integrals.

One mitigating property (that we have yet to make use of) is that for a given fixed direction (κ_s, λ_s), either all points (u_s, v_s) map to the destination quadrilateral or none do.

5.3.2 The Projected Hypercube Has Curved Sides

Even if all points in the source hypercube map to the destination plane, the nature of the resulting volume, is not trivial. Needless to say, the projection of a source parametric hypercube into destination parametric space is not a hypercube. Unfortunately, it is not even a polytope.

Consider the four steps of the transform sequence illustrated in Figure 6. Only the source-object-to-world and world-to-destination-object transforms are affine. The source-parametric-to-source-object transform is bilinear and the the destination-object-to-destination-parametric transform is inverse bilinear[6].

The bilinear source-parametric-to-source-object transform deforms the hypercube into a prism, but once the prism is projected into destination object space, its sides generally do not correspond to surfaces of constant destination parametric coordinates. An arbitrary line of the form $ax_d + by_d = c$ under the inverse bilinear transform contains a term in $u_d v_d$, making the resulting line a hyperbola.

As a result, the projection of the source hypercube into destination parametric space has curved sides. Furthermore, the curvature is such that we cannot guarantee that the convex hull of the polytope formed by projecting the 16 corners of the source hypercube into destination space contains the hypervolume.

5.3.3 Integration Techniques

For these reasons, we must resort to multidimensional pointwise numerical integration schemes. At present, we adopt an unstratified Monte Carlo approach. Other possibilities we are also considering are: stratified Monte Carlo, trapezoidal, and a collection of quasi-Monte Carlo methods as described in Glassner [glas95].

[6] We have to be careful inverting a bilinear transform: For a quadrilateral, points in object space can map to zero, one, or two points in parametric space. We have to determine inclusion in the destination quadrilateral in *object* space first. Then we can reliably invert the coordinates.

5.4 Results

For a test configuration, we imagine light shining through the square stained glass window shown in Figure 7. The incident light shines down at an angle of 45° from the horizontal and is diffused by the glass according to a distribution proportional to the 4th power of the cosine of the angle between the propagation direction and the incident direction. The light falls on the floor, represented by another square.

Figure 8 shows the result, rendered with the help of Craig Kolb's [kolb91] *radshade* raytracer. Due to time constraints, the actual projected image is 16 by 16 pixels (*rayshade* interpolates), but given the diffuse nature of the problem, the resulting image looks adequate and contains none of the "noise" common to the usual Monte Carlo approach to this sort of problem.

5.5 Work in Progress

The wavelet transport solution used for Figure 8 took several hours to propagate a sparse 16x16x16x16 set of wavelet coefficients (in red, green, and blue bands) on an SGI Indigo 2 workstation. Needless to say, this is impractical.

Nevertheless, we are continuing work with improved transport coefficient integration techniques and plan to take further advantage of wavelet representations, such as knowledge of the destination's reflective properties, to reduce the amount of computation required and allow us to go to larger sets of coefficients.

Ultimately, we hope to incorporate wavelet radiative transfer and surface interaction into a global illumination scheme.

References

[chri94] Per Christensen, Eric J. Stollnitz, David H. Salesin, and Tony DeRose. "Global Il-

lumination of Glossy Environments using Wavelets and Importance". Technical Report UW-CSE-94-10-01, University of Washington Department of Computer Science and Engineering, November 1994.

[cohe93] Michael F. Cohen and John R. Wallace. "Radiosity and Realistic Image Synthesis". Academic Press Professional, San Diego, CA, 1993.

[daub92] Ingrid Daubechies. *Ten Lectures on Wavelets*, Vol. 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* SIAM, Philadelphia, PA, 1992.

[fole90] James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes. *Computer Graphics, Principles and Practice, Second Edition.* Addison-Wesley, Reading, Massachusetts, 1990.

[four90] Alain Fournier, Eugene Fiume, Marc Ouellette, and Chuan K. Chee. "Fiat Lux". Technical Report 90-1, Dynamic Graphics Project, University of Toronto, January 1990.

[glas95] Andrew S. Glassner. *Principles of Digital Image Synthesis.* Morgan Kaufmann, 1995.

[gort93] Steven J. Gortler, Peter Schroder, Michael F. Cohen, and Pat Hanrahan. "Wavelet Radiosity". *Computer Graphics Proceedings, Annual Conference Series, 1993*, pp. 221–230, 1993.

[kolb91] Craig E. Kolb. *Rayshade User's Guide and Reference Manual*, release 4.0 edition, January 1991.

[nuss28] W. Nusselt. "Grapische Bestimmung des Winkelverhaltnisses bei der Warmestrahlung". *Zeitschrift des Vereines Deutscher Ingenieure*, Vol. 72, No. 20, pp. 673, 1928.

[schr93] Peter Schroeder, Steven Gortler, Michael Cohen, and Pat Hanrahan. "Wavelet Projections for Radiosity". *Fourth Eurographics Workshop on Rendering*, pp. 105–114, June 1993.

Figure 7: Spatial Component of Test Configuration

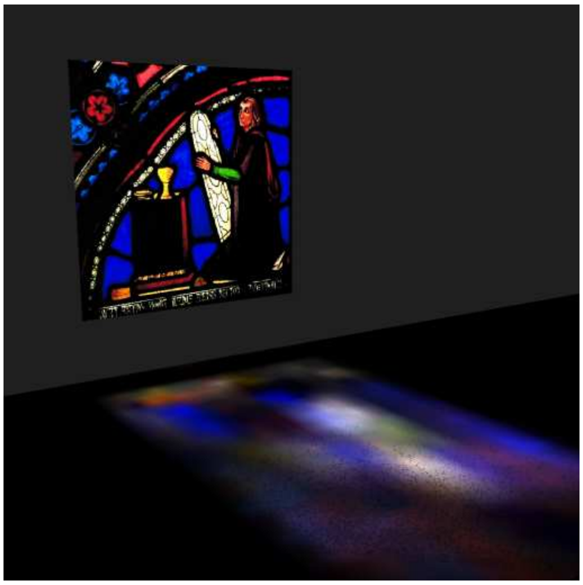

Figure 8: Example of Wavelet Transport