Illumination IV: Radiosity

The whole philosophy of our previous lectures on illumination were based on what we called "quick-and-dirty" methods: efficient approaches that manage to "fool the eye". The *local* illumination is central to the efficiency of these approaches.

A *global* illumination model is a model which take into account the fact that light is not just coming from a few point light source, but that light is arriving indirectly form many different directions.

What are the elements of a global illumination model?

The basic idea is vewing each object as being a potential light source. Some objects (light sources) radiate light directly, but others (nonblack surfaces) can radiate light indirectly.

Radiosity is an example of a global illumination model.

Red

White
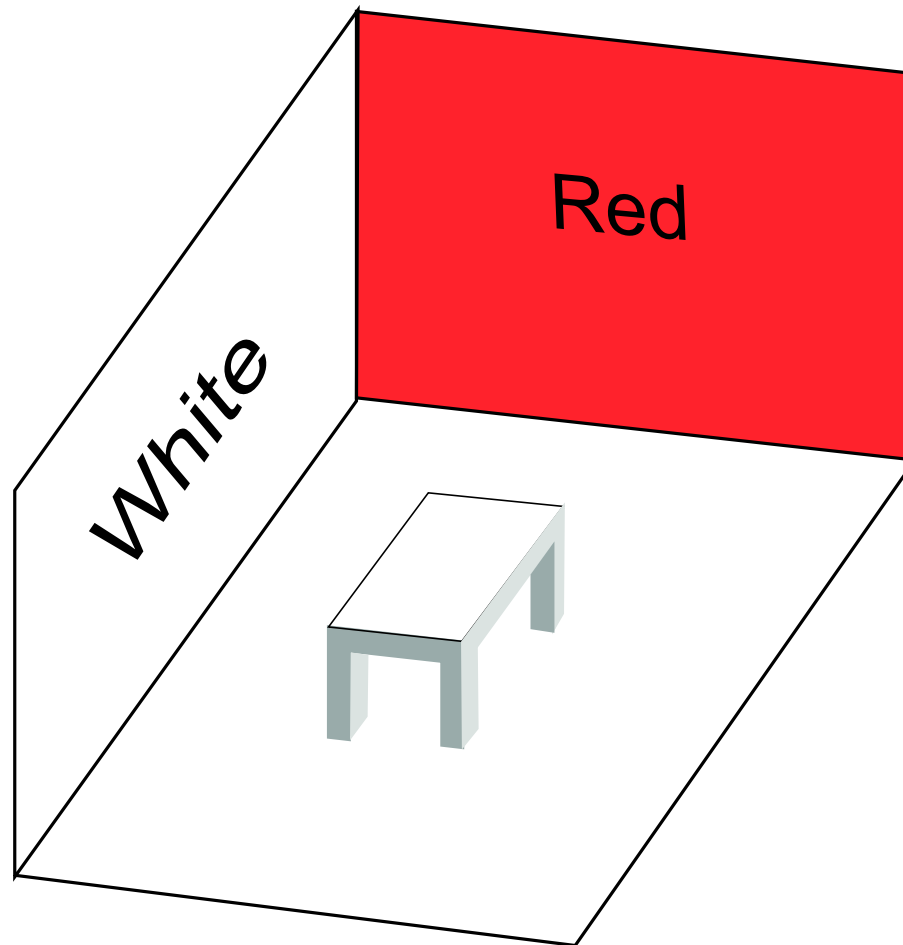
Radiosity Overview

Radiosity: the intensity of each point on the surface of some object in our environment.

This intensity of the point P is a function of

- the emittance of light from this point (if it is a light source),
- the reflection of light coming from other surfaces in the environment.

The second component is quite complicated, because it depends on the radiosity of points on surfaces throughout the environment, whether these points are visible from P, and how reflective the surface is that P lies on.

Sampling

Radiosity computations are quite expensive since for every point we need to know the illumination of all the surface elements that this point can see. A common way is to choose some sampled points in the environment.

How to selected?

The most common way is based on a generalization of the *finite element method.*

- Subdivide each of the object surface into a number of small polygonal patches (*surface mesh*)
- For each patch, compute an approximation of the radiosity of this patch. For example, this could be done by computing the radiosities at each of its vertices and then averaging these.

How to construct these patches?

- Small patches can give good accuracy, but expensive.
- Large patches can give speed, but lost of accuracy.

The best method is to use an *adaptive* approach:

- First start with a coarse mesh, determining in which areas the radiosity is varying most rapidly.
- Then refining these ares and trying again.
- When the radiosity values are fairly constant in the neighborhood of a patch of the mesh, or when the patches are deemed to be "small enough" then we do not need to refine further.
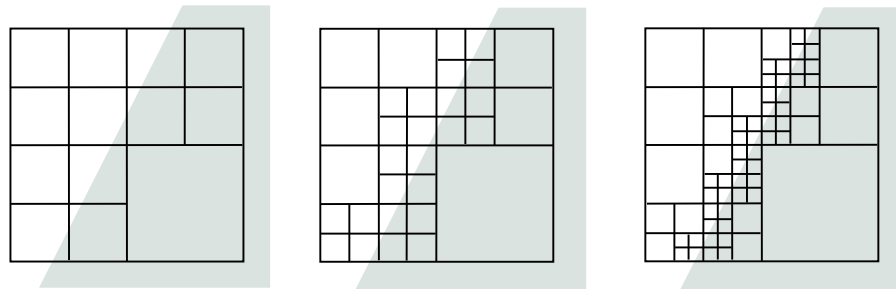
More sophisticated methods, like *discontinuity meshing* actually attempt to align the edges of the mesh with sharp changes in radiosity (e.g. as happens along the edge of a shadow).

Who Comes First?

The radiosity at point A depends on the radiosity from all visible points B. And visa versa. Then how to compute radiosities?

There are two general approaches.

- Define a large linear system of equations, that "encodes" all of the radiosity dependencies. By solving this equation, we can determine all the radiosities at all the points.
  The problem is the size of this linear equation is enormous.
  n surface patches $\Longrightarrow n^2 \times n^2$ matrix.
- *Progressive refinement radiosity:* The idea is:
  - starting with the brightest light source and shooting its radiation around to the entire scene.
  - Then we move to the next brightest light source and repeat this process.
  Note that as we do this, surface that were initially black start picking up more and more intensity. Eventually a nonemitting light source can start accumulating more and more intensity, until it becomes the brightest light source, and then it shoots its intensity to the surrounding scene.

Energy Balance Equation

$$L^o(x, \theta_x^o, \phi_x^o, \lambda^o) = L^e(x, \theta_x^o, \phi_x^o, \lambda^o) +$$

$$\int_0^{\frac{\pi}{2}} \int_0^{2\pi} \int_{\lambda_{\min}}^{\lambda_{\max}} \rho_{bd}(x, \theta_x^i, \phi_x^i, \lambda^i, \theta_x^o, \phi_x^o, \lambda^o)$$

$$\cos(\theta_x^i) L^i(x, \theta_x^i, \phi_x^i, \lambda^i) d\lambda^i \sin(\theta_x^i) d\phi_x^i d\theta_x^i$$

- $L^o(x, \theta_x^o, \phi_x^o, \lambda^o)$ is the radiance
  - at wavelength λ^o
  - leaving point x
  - in direction θ_x^o, ϕ_x^o
- $L^e(x, \theta_x^o, \phi_x^o, \lambda^o)$ is the radiance emitted by the surface from the point
- $L^i(x, \theta_x^i, \phi_x^i, \lambda^i)$ is the incident radiance impinging on the point
- $\rho_{bd}(x, \theta_x^i, \phi_x^i, \lambda^i, \theta_x^o, \phi_x^o, \lambda^o)$ is the BRDF at the point

- – describes the surface's interaction with light at the point
- the integration is over the hemisphere above the point
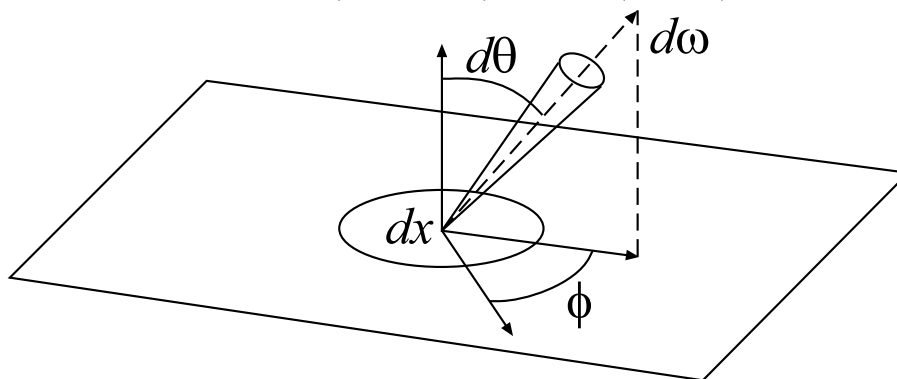
The most basic concept of radiosity is radiance.

Radiance L: the amount of energy per unit time (or equivalently power) emitted from a point x in a given direction.

Define:

- θ: the angle with respect to the surface normal,
- ϕ: the angle of the projection onto the surface.
- ω: the resulting directional vector.

Thus

$$L = L(x, \theta, \phi) = L(x, \omega)$$

The power radiating from a small patch in some small solid angle can be expressed as:

$$L(x, \theta, \phi) dx \cos \theta d\omega$$

Radiance is measured in watts per square meter per steradian.

$$d\omega = (\sin \theta) d\theta d\phi$$

Then the $radiosity$, denoted by B, is as follows

$$B(x) = \int_{\Omega} L(x, \theta, \phi) \cos \theta d\omega$$

where Ω is the hemisphere's surface lying on the above the surface.

Simple Radiosity Equation

If surfaces are Lambertian, then we can simplify $L(x, \theta, \phi)$ and just write $L(x)$. The radiosity at the point x is given by

$$B(x) = \int_\Omega L(x, \theta, \phi) \cos \theta d\omega$$

$$= L(x) \int_\Omega \cos \theta d\omega$$

$$= L(x) \int_0^\pi \int_0^{2\pi} \cos \theta \sin \theta d\theta d\phi$$

$$= \pi L(x).$$

This means simply that depends only on the radiance, the light power, at the point.

Radiosity equation (for Lambertian reflectors):

$$L(x) = L_e(x) + \frac{\rho_d(x)}{\pi} \int_\Omega L_i(x, \theta, \phi) \cos \theta d\omega$$

where L_e denotes emitted radiance and L_i denotes the incoming irradiance, $\rho_d(x)$ denotes the coefficient of diffuse reflection (earlier we had written this k_d).

We cannot eliminate the directional component from the L_i term, because we still need to consider Lambert's law for incoming radiation.

If we define

$$H(x) = \int_\Omega L_i(x, \theta, \phi) \cos \theta d\omega$$

and let $E(x)$ denote the emitted radiosity $\pi L_e(x)$, and recall that $B(x) = \pi L(x)$ then we can write this as

$$B(x) = E(x) + \rho_d(x)H(x)$$

The term $H(x)$ essentially describes how much illumination energy is arriving from all other points in the scene.

To simplify $H(x)$ we can use the Lambertian assumption. Rather than integrating over the angular space surrounding x, instead we will integrate over the set of points on all surface,

denoted S. Let $y \in S$ be such a surface point visible from x in direction ω. Let θ' denote the angle between the surface normal at y and the line-of-sight vector from y to $x(-\omega)$, and let ϕ' be defined similar to ϕ but for y. Let r denote the distance from x to y.
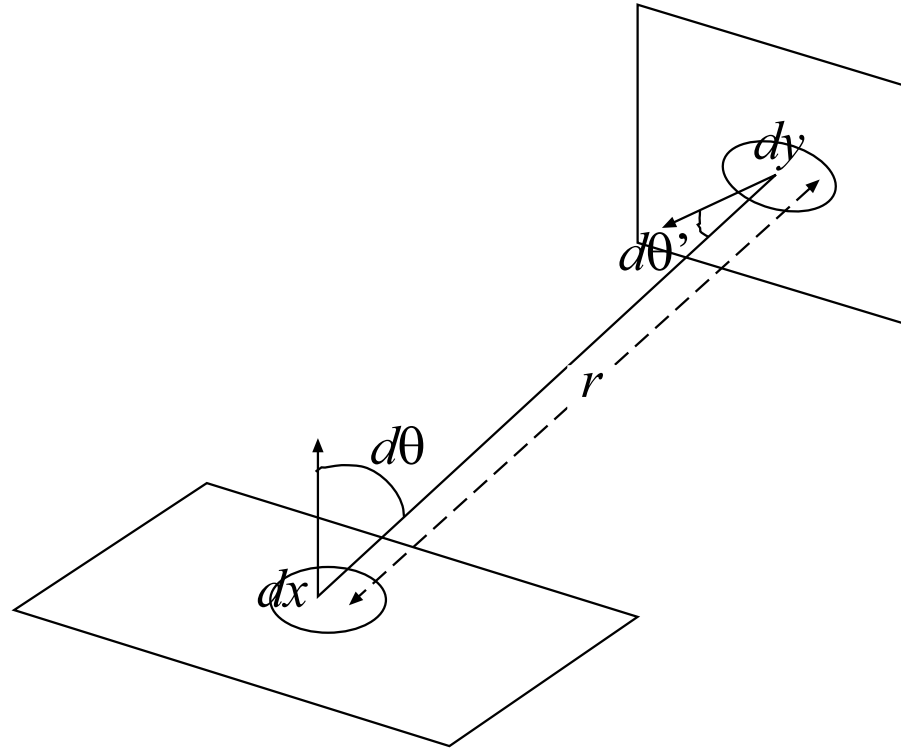
By symmetry of radiance, we have $L(x, \theta, \phi) = L(y, \theta', \phi')$.

Since all surfaces are Lambertian, we have

$$L(y, \theta', \phi') = \frac{B(y)}{\pi}.$$

And

$$d\omega = \frac{\cos \theta' dy}{r^2}.$$

Putting these together, we can define $H(x)$ in terms of an integral over surface points:

$$H(x) = \int_{y \in S} B(y) \frac{\cos \theta \cos \theta'}{\pi r^2} V(x, y) dy.$$

where

$$V(x, y) = \begin{cases} 1, \text{ if } x \text{ can see } y \\ 0, \text{ otherwise} \end{cases}$$

is the *visibility function*.

Form Factors

In practice, we cannot expect to be able to solve this integral equation. As mentioned before, most radiosity methods are based on subdividing space into small patches, and assuming that the radiosity is constant for each path. Thus, in the equation for $H(x)$ above, we can assume that $B(y)$ is constant for all points y in a surface patch.

Form factor $F_{i,j}$: the fraction of light energy leaving P_i that arrives at patch P_j:

$$F_{i,j} = \frac{1}{A_i} \int_{x \in P_i} \int_{y \in P_j} \frac{\cos \theta \cos \theta'}{\pi r^2} V(x, y) dy dx$$

$F_{i,j}$ is a dimensionless quantity. If patches are close, large, and facing one another, $F_{i,j}$ will be large.

Then the radiosity equation is a system of linear equations:

$$B_i = E_i + \rho_i \sum_{j=1}^{n} B_j F_{j,i} \frac{A_j}{A_i}$$

Here B_i is the radiosity of patch i (the amount of light reflected per unit area), E_i is the amount of light emitted from this patch per unit area, ρ_i is the reflectivity of patch i ($\rho \approx 0$ means a dark nonreflecting object and $\rho \approx 1$ means a bright highly reflecting object). A_i and A_j are the areas of patches P_i and P_j, respectively.

The linear system is sparse. Iterative techniques from numerical analysis, such as Gauss-Seidel, can be used to solve this type of system.

Since we assume that light can travel equally well in any direction it follows that

$$A_i F_{i,j} = A_j F_{j,i}.$$

We can simplify the above equation as

$$B_i = E_i + \rho_i \sum_{j=1}^{n} B_j F_{i,j}$$

$$E_i = B_i - \rho_i \sum_{j=1}^{n} B_j F_{i,j}$$

Its matrix form is

$$
\begin{pmatrix}
1 - \rho_1 F_{1,1} & -\rho_1 F_{1,2} & \cdots & -\rho_1 F_{1,n} \\
-\rho_2 F_{2,1} & 1 - \rho_2 F_{2,2} & \cdots & -\rho_2 F_{2,n} \\
\vdots & \vdots & \vdots & \vdots \\
-\rho_n F_{n,1} & \rho_n F_{n,2} & \cdots & 1 - \rho_n F_{n,n}
\end{pmatrix}
\begin{pmatrix}
B_1 \\ B_2 \\ \vdots \\ B_n
\end{pmatrix}
=
\begin{pmatrix}
E_1 \\ E_2 \\ \vdots \\ E_n
\end{pmatrix}
$$

The values ρ_i are dependent on the surface types. The hard thing to compute are the values of $F_{i,j}$.

It can be shown that there is a fairly simple geometric interpretation of $F_{i,j}$.

1. Break the i-th patch into small differential elements.

2. For each element consider a hemisphere surrounding this element, and project patch j onto this hemisphere through its center.

3. Then project this projection orthographically onto the base circle of the hemisphere.

4. The value of $F_{i,j}$ is the area of this projection, divided by the area of the circle.

Thus intuitively patches that occupy a larger field of view contribute more to $F_{i,j}$ and patches that are more nearly orthogonal to the surface contribute more.

Computing this orthogonal projection of a spherical projection is somewhat tricky (considering that it must be repeated for every tiny element of every patch), so it is important of speed this computation up, at the cost of the introduction of approximation errors. We can approximate the hemisphere by a hemicube, and discretize the surface of the hemicube into square (pixel-like) elements. We project all the surrounding patches on to each of the faces of the hemicube. (Note that this is essentially a visible surface elimination task, which can be solved with hardware assistance, e.g. using a z-buffer algorithm.) Each cell of the hemicube is now associated with a patch, and we apply a weighting factor that depends on the square of the hemicube, and sum these up.

Needless to say, this process is extremely computationally intensive. We are basically solving a visible surface determination problem at *every* point on the surface of our objects. Much of the research in radiosity is devoted to mechanisms to save computations, without sacrificing realism.